

An Evaluation of the Fitness Report System for Marine Officers

Adam Clemens • Lauren Malone • Shannon Phillips • Gary Lee

with Cathy Hiatt • Theresa Kimble



DRM-2012-U-001003-Final
July 2012

Photo credit line: Jordan Lane, the son of Maj. Leron Lane, pins the rank of major on his father's collar during a promotion ceremony and luncheon held at Millers Landing on Marine Corps Air Station Cherry Point, N.C., Sept. 4. Jordan, his younger brother Jacob, and Lane's wife Wendy, took part in pinning on Lane's new rank.

Approved for distribution:

July 2012



Anita Hattiangadi
Research Team Leader
Marine Corps Manpower Team

This document represents the best opinion of CNA at the time of issue.
It does not necessarily represent the opinion of the Department of the Navy.

Approved for Public Release; Distribution Unlimited. Specific authority: **N00014-11-D-0323**
Copies of this document can be obtained through the Defense Technical Information Center at www.dtic.mil
or contact CNA Document Control and Distribution Section at 703-824-2123.

Copyright © 2014 CNA

This work was created in the performance of Federal Government Contract Number **N00014-11-D-0323**. Any copyright in this work is subject to the Government's Unlimited Rights license as defined in DFARS 252.227-7013 and/or DFARS 252.227-7014. The reproduction of this work for commercial purposes is strictly prohibited. Nongovernmental users may copy and distribute this document in any medium, either commercially or noncommercially, provided that this copyright notice is reproduced in all copies. Nongovernmental users may not use technical measures to obstruct or control the reading or further copying of the copies they make or distribute. Nongovernmental users may not accept compensation of any manner in exchange for copies. All other rights reserved.

Contents

Executive summary	1
Room for improvement in implementation of FitRep system	1
Training for RSs and ROs	2
Presentation of RO marks to boards.	3
Issues of possible concern	3
Courses of action to consider	4
Introduction	5
FitRep format and process.	6
Organization of this report	10
Have RS and RO marks become inflated since 1999?	11
Have RS marks risen over the last decade?	11
Have RO marks risen over the last decade?.	14
Have specific marks followed different time trends?	15
Are all marks equally informative?	16
Summary	16
How do marks change as the RS or RO gains experience?	19
How do marks change with years of experience?.	20
Does the early formation of an RS profile follow a pattern?	21
Are there “welcome to the grade” and “room to grow” marks?	22
What are the implications of varying the requirements to calculate an RV?	24
Summary	27
Do RS and RO marks each carry important information?	29
How do RS and RO marks correlate over time?	29
Do marks from pre-1999 FitReps predict marks under the new system?	31

Summary	32
How do FitRep marks differ by observable characteristics?	33
Do FitRep marks correlate with other quality measures?	33
TBS third, GCT score, and commissioning source	33
Educational credentials.	34
Do FitRep marks differ across occupational fields?.	36
Do FitRep marks differ by race and ethnicity?	37
Summary	38
Is there evidence of bias?	39
Do the race/ethnicity- and gender-match between the RS and MRO affect marks?.	39
Does occfield-match between the RS, RO, and MRO affect marks?	40
Is there evidence of differences in general perceptions of different occfields?.	42
Summary	45
How do subjective comments correlate with FitRep marks?	47
What are Marine officers taught about the FitRep system?	51
How might boards' view of FitRep results contribute to the boards' confusion?	55
How might the processes for completing and submitting FitReps be improved?	59
Conclusions	61
Successes of the FitRep system.	61
Additional training for RSs, ROs, and boards	61
Clearer and more informative presentation of RO marks	63
Issues for further monitoring and study	63
Appendix A: Blank FitRep Form	65

Appendix B: Statistical regression results	73
Effect of combat FitRep on PARS	73
Welcome-to-the-grade and room-to-grow effects	74
Mutual predictive power of RV and RO mark	75
Quality	77
Education	79
Race and ethnicity	81
Gender match between RS and MRO	82
Race/ethnicity match between RS and MRO	83
Occfield match	83
Promotion recommendation	86
Glossary	87
List of figures	89
List of tables	91

This page intentionally left blank.

Executive summary

The Marine Corps Fitness Report (FitRep) system is the main determinant of an officer's career designation, promotion, and opportunities for command and resident schooling. The current process was implemented in 1999 to address the previous system's perceived shortcomings—chiefly mark inflation. The Director, Manpower Management Division (MM), asked CNA to review whether the system is accomplishing what the Corps intended. She requested that we focus on officers and consider whether the new system is keeping inflation in check, ensuring fairness for all officers, and helping the various boards select the “best and most qualified” officers.

The FitRep form contains administrative data, descriptions of duties/accomplishments in the present billet, 14 dimensions of performance evaluated by the reporting senior (RS), an overall assessment by the reviewing officer (RO), and subjective comments from both the RS and RO. A FitRep average (FRA) of the 14 scores is converted into a relative value (RV) that describes that Marine's position relative to other reports written by the RS on officers in the same paygrade.

In this study, we examined data from all officer FitReps (January 1999 to August 2011), matched with personnel records for Marine Corps officers. We reviewed the FitRep training curriculum and interviewed representatives throughout MM, general officers who have served on promotion and command boards, instructors at The Basic School (TBS) who teach new lieutenants about FitReps, and captains and majors studying at the Expeditionary Warfare School and Command and Staff College, respectively. To the extent possible, we compared insights from these stakeholders with statistical analyses of the data.

Room for improvement in implementation of FitRep system

Overall, we find that the FitRep system is working well. There is no evidence of rampant inflation at an aggregate level. The inclusion of

both RS and RO assessments appears to be helpful and informative. We find that FitRep marks are consistent with other officer quality indicators, and subject matter experts agree that the system usually results in promoting the best and most qualified officers. We do, however, find evidence that there is some room for improvement.

Training for RSs and ROs

To our knowledge, the only formal training officers receive about FitReps is at TBS, before they have any experience with the system. Because TBS students will not immediately act as ROs, they receive no training on RO responsibilities. Training also does not explain how the RV is generated, and gives the false impression that the RV automatically normalizes FRAs into a “bell curve” distribution.

RSs and ROs do not always act in a way that is strictly consistent with the original intent of the Performance Evaluation System (PES) manual. This may be partly because of the lack of formal training. One example is that RSs tend to award “room to grow” marks the first time they evaluate an officer and then award higher FRAs as they continue to evaluate him or her. This is consistent with training they receive at TBS that their FRAs for a particular Marine “*should*” grow over time, and it could be interpreted as instruction not to consider each reporting period separately from all others.

RO marks are intended to have a distribution referred to as the “Christmas tree,” with few marks at the top in order to help boards identify exceptionally qualified Marines. They are also intended to be a relative assessment within a paygrade. However, the actual distribution of RO marks is dramatically different, and officers in higher grades receive higher RO marks on average. We are aware of no formal training on RO responsibilities.

Boards consider subjective comments as well as numerical marks, and their processes for weighing comments are not common knowledge. RSs and ROs feel that greater transparency about what boards are looking for, and how much weight they assign to specific phrases, would improve fairness.

Presentation of RO marks to boards

Board members are provided a “briefing guide” that summarizes a Marine’s career, including a tabulation of RS and RO marks. RO marks are tabulated in a confusing way, with different FitReps receiving different weights depending on how many reports an RO has submitted. A single FitRep could heavily skew the tabulation if the RO who awarded that mark has also marked hundreds of officers in the same paygrade.

In addition, the FitRep form does not allow ROs to specify how closely they observed the Marine reported on (MRO). In many cases, their assessment may be based on infrequent observations or indirect information. If this information were available, greater weight could be placed on RO marks representing more direct observation.

Issues of possible concern

Several findings should continue to be monitored. First, FRAs are becoming less varied and potentially less informative over time. Preventing inflation is irrelevant if every officer receives the same score.

Our analysis shows that black and Hispanic officers receive lower marks on average than white officers, which is not explained by differences in academic credentials, accession source, or General Classification Test (GCT) scores. The gap is primarily driven by differences in TBS standing.

We also consider the interaction between the demographic characteristics of the MRO and the RS. White RSs award *slightly* lower FRAs to black officers, and vice versa—a result that holds even after controlling for such characteristics as family structure, TBS third, and commissioning source. Male and female officers each receive *slightly* higher FRAs on average from RSs of the opposite gender. These are possible indications of gender and racial/ethnic biases or other unexplained variables.

In addition, there may be positive biases toward some occupational fields (occfields) and negative biases toward others. We find that infantry, logistics, and military police officers receive higher marks

than other officers from the same TBS third, and they receive higher marks from ROs in occfields *other* than their own. The reverse is true of aviators. That is, officers in other occfields may be biased toward infantry, logistics, and military police officers and against aviators.

Courses of action to consider

We do not recommend any major overhaul of the FitRep system. We believe, however, that a few minor adjustments may be helpful.

We recommend expanding education and training on FitReps. Officers could be taught how the RV is constructed, and professional military education (PME) for both junior officers and senior enlisted could include insight into the board process, including an online view of the boardroom application and taped demonstrations of boards discussing fictional officers and/or enlisted personnel. MM is developing some training about the board process, which will require cooperation from Training and Education Command.

The briefing guide for board members could tabulate RO marks in a more intuitive way that follows the same logic as its tabulation of RVs. Specifically, for each FitRep that an officer has received with an observed RO mark, the briefing guide could display whether the mark was above, equal to, or below the RO's median for that pay-grade.

Subject matter experts said that ROs often make an assessment of performance based on indirect observation. We suggest that the Marine Corps consider an alternative box for ROs to check that indicates indirect observation so that boards can assign more weight to RO assessments that are based on *direct* observation.

We recommend continued monitoring of the variation in FRAs, of the differences in behavior between white and black RSs and between male and female RSs, and of the possible bias toward infantry, logistics, and military police officers and against aviators. We also recommend further study of the gap between white and minority officers in both TBS performance and FitRep marks.

Introduction

The Fitness Report (FitRep) is an evaluation tool filled out by a Marine's reporting senior (RS) and reviewing officer (RO) that communicates the reporting officials' assessments of the Marine's performance and character to a variety of boards. Today's FitRep system was implemented on January 1, 1999. Like the previous system, it supports promotion boards' selection and retention of the most qualified Marines in the grades of sergeant through major general, the slating of officers for command or resident school billet assignments, and all other billet assignments.

The Director, Manpower Management Division, asked CNA to review whether the system is accomplishing what the Marine Corps intended. Specifically, the director would like to know whether the current FitRep system is suffering from grade inflation, whether it is fair to all officers, and whether it contributes to the challenges that leaders and manpower process managers face in selecting the most qualified officers for promotion and career progression. Although enlisted Marines in paygrades E-5 through E-9 also receive FitReps, we were asked to focus solely on Marine officers.¹

In this study, we address the following questions:

1. Have RS and RO marks become inflated since 1999?
2. How do marks change as the RS or RO gains experience?
3. Do RS and RO marks *each* contain important information?
4. How do FitRep marks differ by observable characteristics, such as quality measures, occfield, gender, and race/ethnicity?

1. We would like to thank our Marine Corps liaison, Capt Nathan Emery, and Maj Christopher Cannon of the Operations Analysis Division for their contributions to this study.

5. Is there any evidence of bias, and, if so, what appear to be the drivers of bias?
6. How do subjective comments correlate with FitRep marks?
7. What are Marine officers taught about the FitRep system?
8. Are FitRep results presented to boards in a clear and intuitive manner?
9. How might the Marine Corps facilitate the process of completing and submitting FitReps?

Before presenting answers based on our analysis, we provide some background information on the FitRep format and process.

FitRep format and process

The RS is the first commissioned officer, warrant officer, or civilian grade GS-9 or above in the reporting chain who is senior to the Marine reported on (MRO). For example, the RS may be either the MRO's commanding officer or the head of a staff section. The RS grades the officer on performance and qualities. Until June 2002, RSs assigned numerical scores to MROs through the rank of colonel. Since then, colonels have not been assigned RS marks.

The RO is the first commissioned officer, warrant officer, or civilian grade GS-10 or above in the reporting chain who is senior in grade to the RS; he or she reviews and decides whether to concur with the RS's report. The RO also ranks the MRO against all Marines in the same grade who are known to the RO. All officers through the rank of colonel receive this relative assessment. FitReps for general officers contain only administrative data and subjective comments.

Marine Corps Order P1610.7F provides detailed instruction on completing and submitting the FitRep.² The current FitRep differs from the one used before 1999 in the following ways:

2. Marine Corps Order P1610.7F W/CH 1. *Performance Evaluation System*. May 2006. Appendix A contains a blank FitRep form.

- The RS grades an MRO on 14 evaluation dimensions referred to as performance anchored rating scales (PARS).³
 - PARS are scored from A (lowest, 1 point) to G (highest, 7 points), with H indicating that the MRO was not observed on that dimension. Any FitRep with even one PARS marked A represents an “adverse” report. The form provides specific descriptions of what merits a B, D, or F within individual PARS, but no description for A, C, E, or G. Marks of A, F, or G must be justified with comments.
 - The 14 PARS belong to 5 sections, as shown in table 1.⁴

Table 1. Performance anchored rating scales

Section	PARS
Mission accomplishment	Performance Proficiency
Individual character	Courage Effectiveness under stress Initiative
Leadership	Leading subordinates Developing subordinates Ensuring well-being of subordinates Setting the example Communication skills
Intellect and wisdom	Professional military education Decision-making ability Judgment
Fulfillment of evaluation responsibilities	Evaluations

- There is no *overall* RS mark; instead, an unweighted average of the PARS is calculated. This is known as the FitRep average (FRA). In the previous FitRep, there was one overall mark for “general value to the service.”

3. Under the previous system, there were 21 rating scales. Only 3 of the current PARS directly correspond to any of the previous 21.

4. These sections did not exist in the previous system.

- The history of all FitReps that an RS has written on Marines in a particular paygrade is tracked and referred to as the RS profile. This profile is used to calculate the MRO's relative value (RV). The RV is a numerical representation—scaled between 80 and 100—of how a single MRO's FitRep compares with other reports written by the same RS on Marines of the same grade.⁵ This pool of comparison includes both unrestricted line officers and limited duty officers of the same paygrade, despite their different responsibilities and experiences. In the previous system, the RS had to provide a relative ranking only if the report was “outstanding” *and* the most recent report on other MROs of the same grade still reporting to the same RS was also “outstanding.” No record was kept of the RS's reporting history.
- The RO provides an overall relative assessment on a scale from 1 to 8, with an intended distribution shaped like a “Christmas tree.” This differs from the RV because it is not derived from other numbers but is a directly assigned relative assessment. There was no numerical RO mark under the old FitRep system.
- There are text boxes for the billet description and accomplishments, whereas there were none in the previous FitRep system. The MRO is responsible for providing a list of his or her own accomplishments to help the RS complete this section. The RS is instructed to encourage the MRO, repeatedly if necessary, to provide this list.
- It is a five-page rather than a two-page evaluation.

5. Specifically, the FitRep with the highest FRA in an RS's profile has an RV of 100, and an FRA equal to the average for the profile would have an RV of 90. The FitRep with the lowest FRA may or may not have an RV of 80, depending on its distance from the average, but the RV cannot be below 80. The formula follows:

$$\max\left(80, \frac{FRA - RS_{average}}{RS_{max} - RS_{average}} \times 10 + 90\right)$$

Under both the current and previous FitReps:

- There is a comment section (the “word picture”) for both the RS and the RO.
- There are boxes for the RO’s concurrence or nonconcurrence with the RS’s marks.

FitReps are required at the following times or for the following reasons:

- Annually for MROs from sergeant through brigadier general in either the active or reserve component (semiannually for second and first lieutenants only)
- Upon the completion of reserve training
- Because of a grade change
- Because of a status change (e.g., deactivation of an active reservist)
- Because of a change of RS (while the MRO remains in the same assignment)
- Because of a change of duty
- Because of a change to (or from) temporary duty
- Because of a transfer (the MRO moves to a new assignment under a new RS)
- If directed by the Commandant of the Marine Corps
- At the end of service

Thus, an MRO receives a FitRep at least once a year. The current FitRep includes administrative information on the MRO, identifies the RS and RO, includes the MRO’s billet description and accomplishments, and displays the 14 PARS to be graded by the RS, the relative assessment of the RO, and the subjective comments of both.

Organization of this report

The remainder of the report follows the order of the numbered list of questions identified at the beginning of this section. Each of the nine numbered items is a separate section. In the first six, we discuss whether there is evidence of grade inflation, systematic unfairness, or inefficiency in the current FitRep system. The next three sections discuss our observations of FitRep instruction, the way FitRep data are presented to boards, and subject matter expert and stakeholder assessments of the submission process. In the final section, we present our recommendations. The appendixes display the FitRep form and contain our statistical regression results.⁶

Following the appendixes is a glossary. For the reader's convenience, we have defined most abbreviations at the beginning of each major section, but the alphabetical glossary is a handy way to look up definitions.

6. We reviewed literature on this topic in November 2011 (see *The Fitness Report System for Marine Officers: Prior Research* by Shannon Phillips and Adam Clemens, CNA Information Memorandum D0026273.A1).

Have RS and RO marks become inflated since 1999?

One of the primary motivations for the current FitRep system was a desire to minimize grade inflation in officers' FitRep scores. Inflation—the average upward trending of *all* FitRep marks over time—weakens the significance of high marks and can make it more difficult to identify the most qualified officers. There is concern that the current FitRep system may again be suffering from inflation. In this section, we look at how the average level and variation of FitRep averages (FRAs) has changed over time within each paygrade, and we examine the average level of reviewing officer (RO) assessments within each grade.

Manpower Management Support Branch (MMSB) provided us with data from the Performance Evaluation System (PES) for all officers since the introduction of the new FitRep system in 1999, including reporting senior (RS) and RO marks, the time period covered, and the identities of the Marines reported on (MROs), RSs, and ROs.⁷ The 14 performance anchored rating scales (PARS) are averaged into 1 FRA for each FitRep submitted.

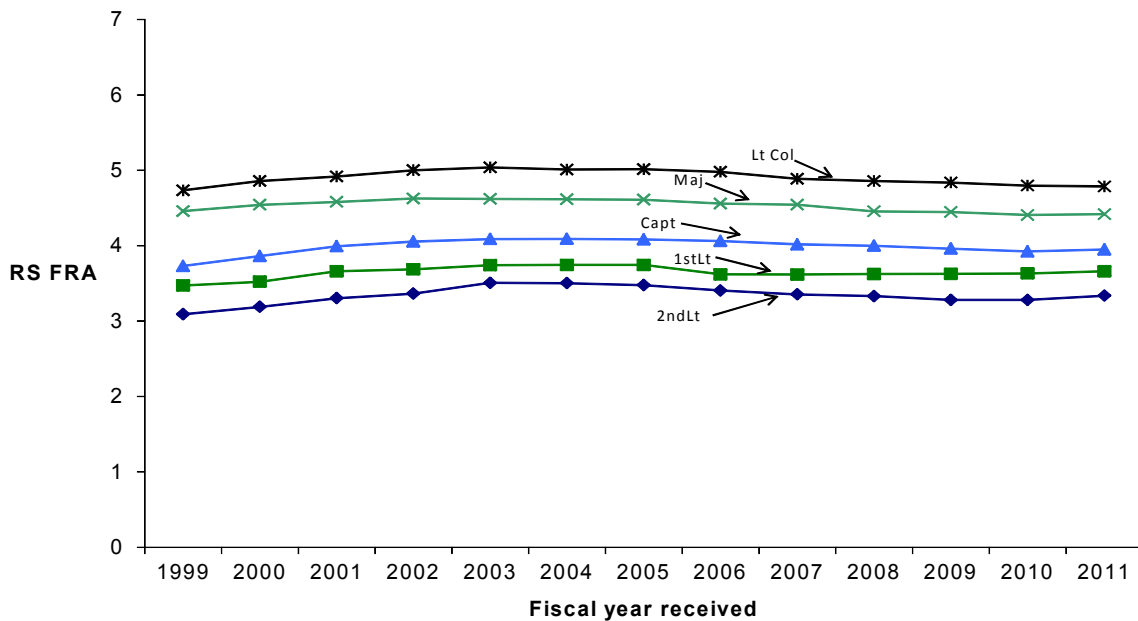
Have RS marks risen over the last decade?

Figure 1 shows average FRAs by fiscal year and paygrade (O-1 through O-5). It is not surprising, given their greater experience, that officers in higher ranks receive higher marks. Within each rank, however, we also find that average marks rose by about 0.4 or less (on a scale from 1 to 7) between FY 1999 and FY 2003. They fell between FY 2004 and FY 2008 and have been mostly level since then. FRAs for captains, first

7. We thank Ms. Doreen Marucci, Ms. Colaine Minor, and Mr. Daniel Holsinger for their assistance.

lieutenants, and second lieutenants were slightly higher in FY 2011 than they were in FY 1999, but lower than they were from FY 2003 through FY 2005.⁸

Figure 1. RS FitRep averages (FRA) by rank of MRO and fiscal year^a



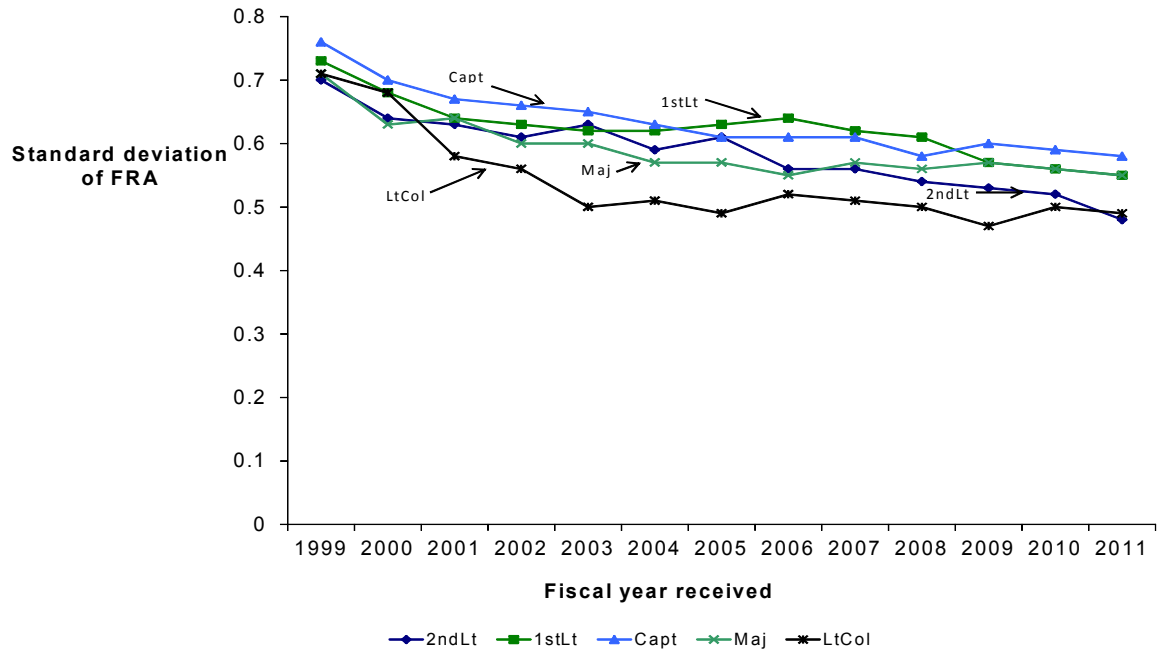
a. Source: CNA analysis of Personnel Evaluation System data from MMSB.

Average levels, of course, are not the whole story. If all captains got an FRA of 4, the average level would look exactly the same as if a quarter of them got a 3, half got a 4, and a quarter got a 5. However, the latter scenario would be more helpful to boards. Therefore, we also look at the standard deviation of FRAs. Figure 2 shows that marks are becoming more uniform over time: the standard deviation of FRAs has

8. Weighting marks according to the length of time reported on does not substantially change these results.

fallen for every rank. This increasing uniformity could increase the difficulty of identifying the most qualified officers.

Figure 2. Standard deviation of FitRep averages (FRA) by rank of MRO and fiscal year^a



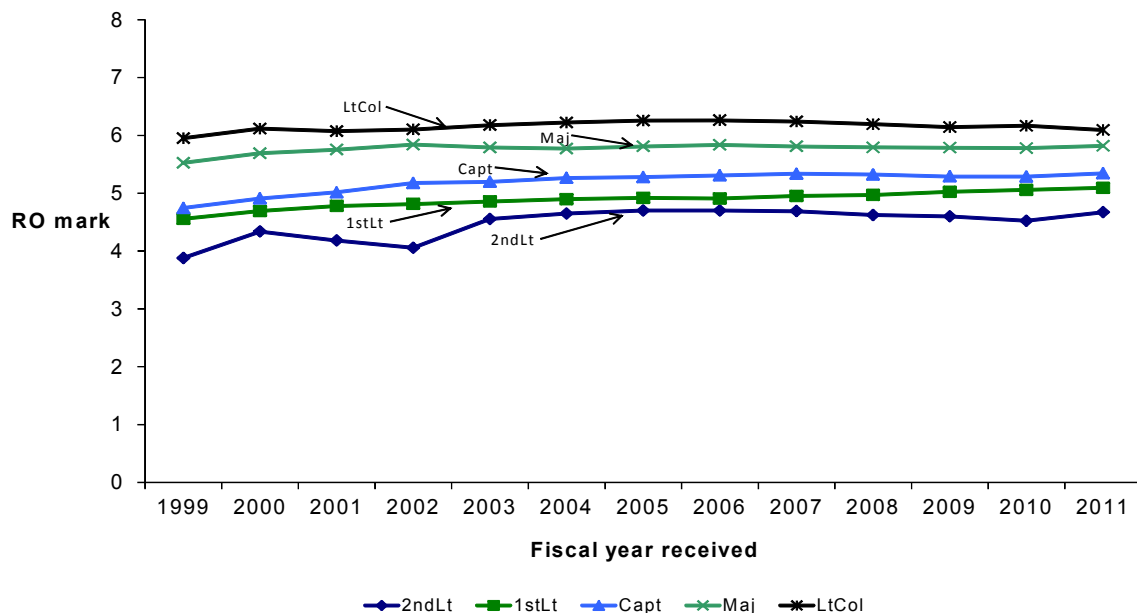
a. Source: CNA analysis of Personnel Evaluation System data from MMSB.

We also find that, for a given MRO grade, higher ranking RSs tend to give higher marks. For instance, captains receive higher marks from lieutenant colonels than from majors, higher still from colonels, and even higher from brigadier generals. This may be partly because of differences in quality between captains who report to majors and captains who report to colonels. It is also possible that the farther the RS is from the MRO's grade, the less visibility the RS has on the MRO's performance.

Have RO marks risen over the last decade?

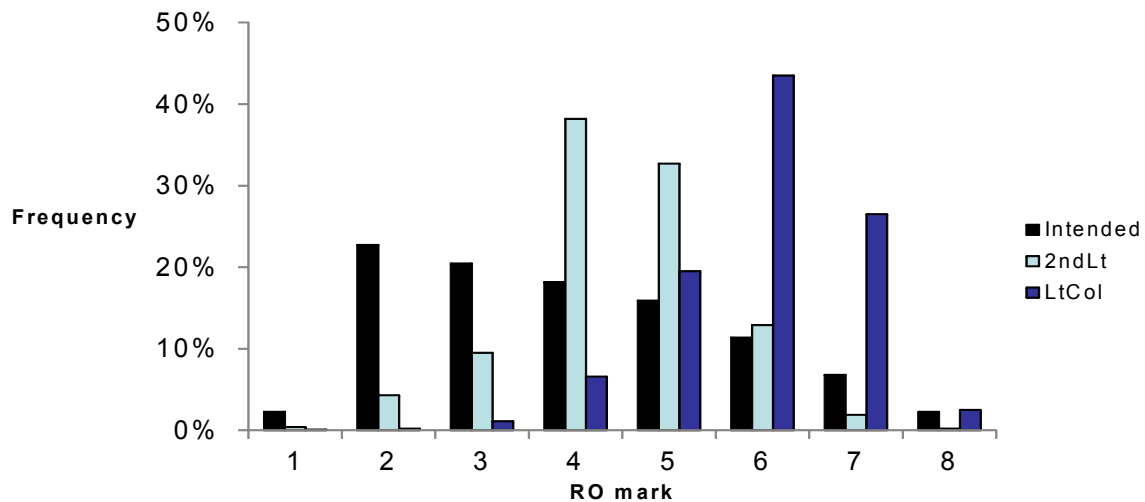
Figure 3 shows average RO marks for each fiscal year and paygrade. Once again, higher ranks receive higher marks, despite the fact that the RO mark is intended to be a comparative assessment relative to other officers of the same grade. The average mark for lieutenant colonels is 6, which is meant to denote “one of the few exceptionally qualified Marines.” Figure 4 confirms that RO marks do not fit the intended distribution, as shown in the Christmas-tree pattern in section K of the FitRep form, suggesting that they are less informative about the true spread of quality than they are intended to be. The distributions for first lieutenants, captains, and majors, which are not shown, lie between those for second lieutenants and lieutenant colonels.

Figure 3. Average RO marks by rank of MRO and fiscal year^a



a. Source: CNA analysis of Personnel Evaluation System data from MMSB.

Figure 4. Distribution of RO marks among second lieutenant and lieutenant colonel FitReps compared with the intended RO mark distribution within a grade^a



a. Source: CNA analysis of Personnel Evaluation System data from MMSB.

RO marks have risen slightly over time. Consider the following changes from FY 1999 to FY 2011: on a scale from 1 to 8, RO marks received by lieutenant colonels, majors, captains, first lieutenants, and second lieutenants were higher by 0.1, 0.3, 0.6, 0.5, and 0.8 point, respectively. While not shown, the standard deviation of RO marks has declined over time, though it has rebounded some since FY 2008. However, we view the changes over time in the level and standard deviation of these marks to be of less concern than the persistently high averages of RO marks in the field grades.

Have specific marks followed different time trends?

Though marks on all PARS followed the same increasing trend from FY 1999 to FY 2003 and the same decreasing trend thereafter, marks on a few PARS had relatively larger increases from FY 1999 to FY 2003. These PARS, which are of particular importance in combat, include mission performance, courage, and effectiveness under stress; they saw increases of 0.40 to 0.42 point. Marine officers who had a FitRep in combat between FY 1999 and FY 2003 scored 0.2, 0.5, and 0.3 point

higher on these three PARS, respectively, than Marines without a combat FitRep.⁹ Combat FitReps had a significant but smaller effect on the remaining PARS. Therefore, Operation Enduring Freedom and Operation Iraqi Freedom may have contributed to the initial increase in FRAs.

PARS whose marks increased the least from FY 1999 to FY 2003 include two leadership PARS (setting the example (0.25 point) and ensuring the well-being of subordinates (0.27 point)), PME completion (0.23 point), and accuracy and timeliness of evaluations (0.12 point).

Are all marks equally informative?

Captains attending the Expeditionary Warfare School (EWS), monitors at the Officer Assignment Branch (MMOA), and former board members all believe that there should be fewer PARS. They discussed disadvantages associated with having 14 different attributes. RSs know that the average value of these 14 marks—as well as the resulting RV—carries far more weight than any individual mark, so they are likely to adjust individual PARS to generate the average they believe is appropriate.

Because of these recommendations, we inspected both where there are redundancies in PARS and where there is little variation—and therefore little value added—in PARS. The most redundant of the PARS is decision-making ability, which is highly correlated with judgment. The PARS with the least variation are fulfillment of evaluation responsibilities and ensuring the well-being of subordinates. However, we believe that improved training could increase the variation and usefulness of both these marks.

Summary

Overall, we find that FRAs rose slightly through FY 2003—partly driven by combat-oriented PARS—but have since declined. They

9. Table 10 in appendix B contains our regression results.

have continually become less varied and, therefore, potentially less informative. RO marks have risen over time, particularly for lower paygrades, but the greater anomaly is that they are higher on average for higher paygrades despite being described as a relative assessment that is separate for each grade.

There is no evidence that grade inflation among officers is rampant at an aggregate level across the Marine Corps. The disparity in RO marks for different ranks, however, suggests that ROs are interpreting their marks differently from the way the FitRep form officially describes them. The declining variation in marks is also a potential cause for concern.

This page intentionally left blank.

How do marks change as the RS or RO gains experience?

As reporting seniors (RSs) submit more FitReps on Marine officers in a particular paygrade, we suspect that their marks may change in a systematic way. There could be at least two reasons for this: (1) Experience in supervising officers of this rank could affect how their relative performance is perceived and evaluated, and (2) concerns about the effect of marks on the RS's profile may change as the profile matures. FitRep averages (FRAs) are converted into a relative value (RV) that is a function of the RS's marks on Marine officers of the same paygrade. Boards make greater use of the RV than the FRA because the former accounts for differences in marking philosophy to enable fairer comparisons between FitReps.

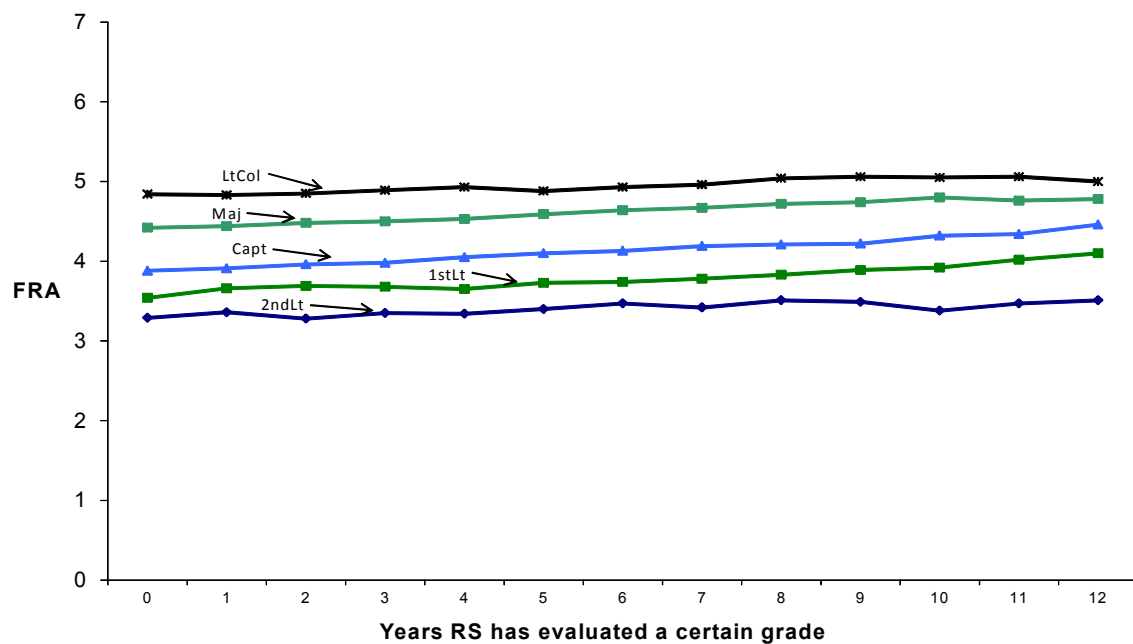
Similar effects could apply to reviewing officers (ROs), although to a lesser extent. When an officer begins evaluating Marine officers of a particular paygrade as an RO, he or she already has experience evaluating that paygrade as an RS, and RO marks are not converted into RVs in the same explicit way that FRAs are.

In this section, we first show the overall trend in FRAs as a function of an RS's experience with a paygrade. Next, we focus specifically on patterns in the first marks that an RS awards to a given paygrade. Neither of these approaches, however, accounts for differences in the Marine reported on (MRO), so—as a last step—we look at patterns of RS behavior from the MRO's perspective while accounting for differences between MROs. Finally, because of concerns about the early stages of an RS profile, stemming both from our analysis here and from interviews with officer career counselors, we consider the implications of adopting more stringent requirements to generate an RV so that fewer FitReps have an associated RV.

How do marks change with years of experience?

On average, the longer an RS has been evaluating Marine officers of a given paygrade, the higher the marks he or she gives MROs in that grade. This effect is more pronounced and persistent for FRAs given to first lieutenants, captains, and majors (see figure 5); it is smallest for lieutenant colonels. We also find that the standard deviation of FRAs falls during the first two years of evaluating experience, but the pattern is inconsistent after that.

Figure 5. Average FitRep marks (FRA), by grade of MRO and years the RS has evaluated that grade^a



a. Source: CNA analysis of Personnel Evaluation System data from MMSB.

We saw in the last section that higher ranking RSs award higher marks. Because higher ranking RSs are also more experienced, we want to separate the effect of evaluating experience from the effect of

paygrade. To do this, we predict the FitRep average as a function of the MRO's grade, the RS's grade, and the number of years the RS has been evaluating that paygrade. We find that one-third of this increase in marks over time is directly a result of experience as an RS. That is, there is a direct effect of evaluation experience, but it is less important than the RS's overall seniority.

Does the early formation of an RS profile follow a pattern?

Beginning RS profiles do not look the same as mature ones. If they did, there would be a constant probability that the next FRA is higher (or the same, or lower) than the last one. That is, in a steady state, the expected relationship between any two FitReps that an RS writes would be independent of the RS's experience. Instead, we find that in all paygrades there is an initial tendency for marks to rise—a tendency that diminishes as the profile matures. Specifically, the probability of the next FitRep having a higher FRA than the previous one diminishes over time, as the profile levels off from an initial climb.

This tendency to climb more steeply at the beginning of a profile holds whether we consider all FitReps in an RS's profile, the FitReps on a particular MRO, or the first FitRep for each of a series of MROs. When we examine FRAs for a particular MRO, we find that an MRO tends to receive higher marks the longer he or she is evaluated by the same RS, although for lieutenant colonels this peaks with the third FitRep.

This tendency of marks to rise as the profile is established works differently for different paygrades, as shown in table 2. For second lieutenants, the RS's marks on a particular MRO rise over time, but the starting point for each new MRO does not increase over time. For first lieutenants, not only do marks rise for a particular MRO, but later MROs start higher than earlier ones. Among captains, the initial climb tends to level off with the fourth overall FitRep or the second MRO, but the standard deviation continues to fall (the marks become more clustered).

Majors and lieutenant colonels run into an apparent "ceiling": the RS may award an FRA slightly above 5 but is not likely to go any higher

because any individual PARS higher than an “E” require justification. Therefore, as the lower range of what the RS might mark drifts up over time, the spread decreases dramatically. An RS who has written on a few lieutenant colonels already appears hesitant to give one less than a 4.5, and will have difficulty justifying more than a 5.2.

Table 2. Median FRA awarded by RS on each successive FitRep he or she writes for officers of a particular paygrade^a

	1 st	2 nd	3 rd	4 th	5 th	6 th	7 th
Order of FitRep in grade							
Second lieutenant	3.23	3.29	3.29	3.31	3.31	3.31	3.36
First lieutenant	3.50	3.57	3.57	3.64	3.64	3.71	3.71
Captain	3.85	3.93	3.93	4.00	4.00	4.00	4.00
Major	4.15	4.36	4.43	4.43	4.43	4.46	4.5
Lieutenant colonel	4.31	4.73	4.86	4.86	4.86	4.86	4.92
Order of MRO in grade (first FitRep for this MRO)							
Second lieutenant	3.23	3.23	3.21	3.21	3.23	3.23	3.23
First lieutenant	3.50	3.50	3.57	3.57	3.62	3.62	3.64
Captain	3.85	3.92	3.93	3.93	3.93	3.93	3.93
Major	4.15	4.31	4.38	4.42	4.43	4.43	4.43
Lieutenant colonel	4.31	4.71	4.85	4.86	4.86	4.86	4.86

a. Source: CNA analysis of Personnel Evaluation System data from MMSB

Are there “welcome to the grade” and “room to grow” marks?

RS and RO marks differ significantly from average for:

- the first FitRep that an MRO receives in a new paygrade (the “welcome to the grade” effect),
- the first FitRep that an MRO receives from a new RS or RO in a given paygrade (the “room to grow” effect), or
- the first FitRep that an RS or RO writes for a new grade of MRO.

These effects, however, differ by paygrade and between RSs and ROs. In this subsection, we predict FRAs and RO marks while controlling for unobserved differences between MROs, months of commissioned service, and calendar year in order to determine where these effects are most likely to occur and how large they are.¹⁰

At all paygrades, both RSs and ROs seem to give room-to-grow marks: an MRO receives a lower mark on his or her first report from a new RS or RO.¹¹ The size of this effect varies between 0.1 and 0.2 point (e.g., an officer receives an FRA of 3.6 instead of 3.8 and has about a 20-percent probability of receiving an RO mark of 4 instead of 5). This room-to-grow effect is about twice as large if the first report covers less than 180 days than if it covers more, but it applies to reporting periods of all lengths.

On top of this, RSs seem to award welcome-to-the-grade marks for new second lieutenants, first lieutenants, and captains. ROs seem to award welcome-to-the-field-grades marks to majors—though the average effect is small, about 0.05 point. That is, a new major has about a 5-percent chance of receiving an RO mark one notch lower than it would otherwise have been.

We note that a change in RS often coincides with a change in responsibilities. It is likely that room-to-grow marks are largely the result of the MROs adapting to new billets or new responsibilities within the same billets, and then performing better as they continue to serve in these responsibilities.

However, room-to-grow marks may also be partly a result of training that officers receive at The Basic School (TBS). TBS instructors suggest that an MRO's marks "should be" increasing over time with the same RS, though there will be exceptions. They do not clearly distinguish between the MRO's performance truly improving over time, while holding the standard of evaluation constant, and raising the

10. Because we do not report the effects of specific years in these regression models and are only interested in controlling for them while studying other effects, the choice of how to divide 12-month periods is arbitrary.

11. Tables 11 and 12 in appendix B contain our regression results.

marks with no apparent change in performance. A desire to do this could lead an RS to award an artificially low FRA the first time.

Being the recipient of the first FitRep an RS has written for a first lieutenant or a captain results in a lower mark, perhaps because the RS is trying to be conservative when beginning a new profile. In contrast, the first FitRep that an RO writes for a new paygrade tends to be higher than otherwise expected. Surprisingly, the first FitRep that an RS writes on a second lieutenant also tends to be higher than otherwise expected. As noted in table 2, second lieutenants are unique because an RS's starting point for each new second lieutenant does not rise over time.

As shown earlier in figure 1, average FitRep marks rose somewhat during the first few years of the new system, but that trend has reversed since at least FY 2005. RSs and ROs are likely to have become more familiar and comfortable with the new system over time. When we split the sample into 1999 through 2005 and 2006 through 2011, we find that the welcome-to-the-grade effect is more pronounced in the more recent sample for both FRAs and RO marks. The room-to-grow effect has also become more pronounced, but only for RO marks. These recent trends may be cause for concern.

What are the implications of varying the requirements to calculate an RV?

FRAs are a problematic tool for comparing Marines because different RSs have different marking philosophies. Therefore, board members are aided by a tabulation of an officer's RVs, which compare their performance against peers in the same paygrade evaluated by the same RS. FitReps written early in the development of an RS's profile can be especially problematic for boards in the near term, when there is a limited pool of peers for comparison.

Currently, an RV is generated for an observed FitRep as soon as the RS has written at least three observed FitReps for that paygrade. From discussions with career counselors, several suggestions arose for tracking and processing FitReps. Because boards view a tabulation of all the RVs an officer has received and, in general, can use the RV as a

convenient tool to help interpret a FitRep, more stringent requirements regarding the generation of an RV would prevent some FitReps from affecting the overall tabulation and would require boards to treat these FitReps differently. Specific possibilities include the following:

1. Raise the minimum number of reports to form an RS profile from three to four or five—requiring more scope for comparison before a FitRep is tabulated with the others.
2. Require a minimum number of different MROs for RV calculation, rather than a minimum number of observed FitReps—thus ensuring that the FitRep is not simply being compared with one peer whom the RS has reported on multiple times.
3. Have limited duty officers (LDOs) compete against one another and, separately, unrestricted officers compete against one another—avoiding an unintended comparison of officers with very different experiences and duties.

For each of the three suggestions, we calculate the percentage of FitReps from which an RV would be generated by the time the MRO either was promoted out of that grade or separated.¹² This serves as an approximation of the percentage with an RV in time for the promotion board. For example, what percentage of FitReps will come from an RS who has written at least three FitReps (or on at least three different MROs, LDOs, or unrestricted officers) in that paygrade *before* the MRO is up for promotion? If a FitRep does not have an associated RV, the board must rely more on subjective comments and context to interpret it.

1. If the number of reports needed to form an RS profile is raised from three to four, the percentage of observed FitReps with an RV falls from 84 to 77 percent (see the first set of columns in table 3). If five FitReps are needed, the percentage falls to 69 percent. The percentage of FitReps with an RV by the time the

12. For all FitReps written on second lieutenants and first lieutenants, we consider when they are granted career designation and promoted to captain or when they separate.

MRO reaches the promotion board in that paygrade is highest for captains, roughly equivalent for lieutenants and majors, and lowest for lieutenant colonels.

2. If we require a minimum number of different MROs in the RS's profile, the percentage of FitReps with an RV in time for promotion falls to 79 percent, 71 percent, or 64 percent, respectively, for three, four, or five MROs (see the second set of three columns in table 3).
3. If LDOs compete only with other LDOs, the share of FitReps for LDOs with an RV falls to 23 percent or less. In most cases, boards would not have the benefit of RVs to evaluate LDOs' performance.
4. If unrestricted officers compete only with other unrestricted officers, the percentages of their observed FitReps with an RV are virtually identical to those appearing in the right-hand side of table 3; this change would come at very little cost from the perspective of unrestricted officers.

Table 3. Percentage of FitReps with an RV generated by the time the MRO is promoted, by number of reports and by number of different MROs needed to form RS profile

	Number of FitReps needed			Number of different MROs needed		
	Three	Four	Five	Three	Four	Five
Career designation	0.84	0.76	0.68	0.76	0.66	0.58
Captain	0.86	0.79	0.72	0.83	0.75	0.69
Major	0.83	0.75	0.68	0.79	0.71	0.63
Lieutenant colonel	0.81	0.74	0.68	0.77	0.70	0.64
Average	0.84	0.77	0.69	0.79	0.71	0.64

Career counselors observed that the RV currently compares a FitRep with all others in the RS's profile for that grade, including others written on the same MRO. Therefore, an MRO competes against himself or herself, and a high mark will lower the RV of his or her other reports from that RS. There is no reason why an officer should be penalized for his or her own strong performance, or rewarded for his or her own weak performance.

Summary

Higher ranking RSs award higher marks to a given rank of MRO, and RSs award higher marks the longer they have been evaluating Marine officers of a given rank. There is also evidence that RSs have a systematic approach to establishing their profiles because the first few FitReps they write tend to have climbing FRAs before gradually settling into a more stable pattern. For example, the first FitRep RSs write for a lieutenant colonel has an FRA more than half a letter grade lower, on average, than the seventh FitRep they write for that paygrade. These effects may partly be unintended consequences of the relative value system and the instruction officers receive about FitReps.

There is also evidence that officers receive welcome-to-the-grade and room-to-grow marks from RSs and ROs, although the size of these effects is less than a fifth of a letter grade on average. This leaves boards with the responsibility of guessing how actual performance would have been reported in the absence of these trends—which they are probably not able to do perfectly—and may result in unfairness.

This page intentionally left blank.

Do RS and RO marks each carry important information?

As discussed earlier, the reporting senior (RS) evaluates 14 characteristics that are averaged into 1 number (the FRA) and converted into a relative value (RV). The reviewing officer (RO) makes 1 relative assessment. Do these assessments move together, or do they differ? If they differ, why?

How do RS and RO marks correlate over time?

To compare RS marks (indicators of absolute performance) with RO marks (indicators of potential relative to other officers within a paygrade), we use the RV rather than the raw FRA. The RO mark is a relative assessment, so we compare it with another relative assessment. Table 4 shows the spread of RVs corresponding to each RO mark for captains. In all paygrades, there is a strong correlation between the RV and the RO assessment on the same report. For instance, the median RV for a captain with an RO mark of 2 is 80, and each successively higher RO mark is associated with a higher median RV: 3 with 82.9, 4 with 87.8, 5 with 91.2, 6 with 95.6, and 7 or 8 with 100.

However, the RV and RO mark do not move in lock step. There is also considerable overlap within each grade: an RO may award a 4 when the RS awards marks generating a high RV, or an RO may award a 6 when the RV is low.

There are at least three possible explanations for why RVs and RO marks may differ:

1. One of these two values is the “true” assessment of performance; any difference from the other is random and has no value for predicting future performance; this is a “straw man,” which we hypothesize in order to disprove.

2. The RS and the RO make partially independent assessments of the same qualities; each assessment has value, but neither is a perfect measure on its own.
3. RSs and ROs are looking for different things, or have different biases, perhaps based on different experiences.

Table 4. Distribution of RVs for FitReps with a particular RO mark among MRO captains, January 1999 to September 2011^a

RV distribution	RO mark							
	1	2	3	4	5	6	7	8
Number of FitReps	198	312	2,044	12,467	23,487	18,061	4,636	266
10 th percentile	80	80	80	80	83.2	87.5	91	92.3
25 th percentile	80	80	80	83.7	87.5	91.3	94.9	96.8
Median (50 th percentile)	80	80	82.9	87.8	91.2	95.6	100	100
75 th percentile	80	83.2	87	91.4	95.9	100	100	100
90 th percentile	83.1	87.5	90.5	96.7	100	100	100	100

a. Source: CNA analysis of Personnel Evaluation System data from MMSB.

Empirically, we find support for both explanations 2 and 3. Both the RV and the RO mark become more informative as they are averaged over an officer's career. For instance, the average of RVs over a career and the average of RO marks over a career are more correlated than the RV and RO mark on a single FitRep. Our best test of informative value, though, is how well the past can predict the future. We can make a prediction of the current report's RV or RO mark using just the value of the other assessment on this report, or using just the past history of the same assessment.

On one hand, both past RO marks and the current RV help to predict the current RO mark, but we can improve the accuracy of either prediction by including the past history of the other assessment (e.g., past RVs also help to predict the current FitRep's RO mark).¹³ On the other hand, we find that the RV and the RO assessment systematically

13. Table 13 in appendix B contains our regression results.

differ for a given Marine reported on (MRO). If Major Brown has a history of earning higher RO marks relative to his RVs, we predict that his RV on this FitRep will once again be low relative to the RO assessment. This is true in all paygrades, and it remains true when we control for year, grade, gender, race or ethnicity, occupational field (occfield), TBS third, commissioning source, marital status, and/or General Classification Test (GCT) score.

Therefore, there is evidence that the RS and RO are, to some extent, assessing the same thing (since they each help to predict the other), and to some extent assessing different things (since they differ systematically for a particular MRO). Each of the next two sections contains examples of how the perspectives of the RS and RO may differ. Before moving on, we consider whether RSs are also to some extent assessing the same things that they were assessing before 1999.

Do marks from pre-1999 FitReps predict marks under the new system?

We observed a sample of FitRep marks for officers between CY 1996 and CY 1998.¹⁴ A little over 8,000 officers had FitReps both in our pre-1999 sample and in the current system. In addition to the FRA being analogous to the “general value to the service” mark, three attributes are nearly identical in the old and new FitReps: initiative, judgment, and leadership (“leading subordinates” on the new form). If FitReps were accurately measuring performance both before and after the transition, high marks on these attributes before 1999 should predict high marks in the new system as well.

In all paygrades, the FRA is positively correlated with general value to the service, and initiative, leadership, and judgment are each positively correlated with their pre-1999 counterparts.¹⁵ FRAs were 0.48

14. This is a smaller sample window than previously presented because the priority of this study is to evaluate the new system, not the old.

15. This is true whether we compare (a) an officer’s first FitRep under the new system with the last under the old system or (b) the average of all the officer’s reports after the transition with all those we observe before.

point higher in the current system for officers whose general value had been consistently rated “outstanding” before 1999 than for officers who had at least one general-value mark lower than outstanding between 1996 and 1999. This gap is even larger for officers who were of higher rank at the time of the transition: a mark of less than outstanding may have been increasingly rare, and therefore increasingly informative, in higher paygrades.

We note that 89 percent of these officers were rated outstanding in every FitRep they received between 1996 and 1998—higher than the percentage that can be promoted to major. Thus, it is clear that FitRep marks are more informative under the new system.

Initiative, leadership, and judgment marks since 1999 are each more strongly correlated with the pre-1999 general-value mark than with pre-1999 marks for initiative, leadership, or judgment. This may suggest that these individual marks are not interpreted the same way now as they were in the previous system, or that the general-value mark—graded on a 10-point scale—was slightly more informative than the pre-1999 individual marks that were graded on a 6-point scale.

Officers who were rated lower in the final years of the previous FitRep system continued to be rated lower under the new system, suggesting that both systems were assessing some proxy for true quality. This effect is strongest for the old general-value mark and the new FRA.

Summary

RS and RO marks each provide valuable information. They are, to some extent, measuring the same things because each helps to predict the other. The RS average also, to some extent, measures the same things as the previous general-value-to-the-service mark. However, the RO and RS marks also measure different things: there is a tendency for some officers to consistently receive more favorable (or unfavorable) evaluations from ROs than RSs. The next two sections include examples of how RS and RO perspectives may differ.

How do FitRep marks differ by observable characteristics?

Next, we examine how FitRep marks vary with personal characteristics, such as commissioning source, GCT score, occupation, and demographics. We also look for clues about what forms of bias may still be present in the FitRep system. To do this, we merge PES data with personnel data from CNA's Headquarters Marine Corps Master File, extracted from the Total Force Data Warehouse (TFDW). In this section, we describe the effect of general quality indicators on FitRep marks and show how marks differ across occfields, race/ethnicity, and gender. The next section will contain a more detailed examination of possible drivers of these differences.

Do FitRep marks correlate with other quality measures?

To the extent that FitReps accurately reflect officers' performance and potential, we would expect them to correlate with other observable characteristics. First, we look at the effect of military-specific quality measures on FitRep marks. Then, we examine the effect of educational credentials.

TBS third, GCT score, and commissioning source

We investigate how observable measures of quality—TBS ranking (a composite ranking made up of academic, leadership, and military skills), GCT score, or commissioning source—predict FitRep averages (FRAs). We find that officers ranked in the top third of their TBS classes have higher FRAs than those ranked in the middle third, and those in the bottom third have still lower marks. This holds controlling for year, grade of the Marine reported on (MRO), gender, race/ethnicity, marital status, and dependents. It is also consistent with previous studies, as noted in our separately published review of prior research. The boost to FRA from being in the top relative to the

middle third or the middle relative to the bottom third is larger for officers in lower paygrades; it diminishes as they pass competitive hurdles (e.g., career designation, promotion to lieutenant colonel).¹⁶

Marine officers with higher GCT scores have slightly *lower* FRAs when we control for race and ethnicity (on average, white officers have higher GCT scores and higher FRAs). This may suggest that the predictive value of the GCT is limited.

Male Marine officers' FRAs differ by commissioning source, while female officers' FRAs do not. Men accessed through enlisted-to-officer (E-to-O) programs¹⁷ score highest, followed by those from the Naval Reserve Officers Training Corps (NROTC), other sources, Officer Candidates Course (OCC), United States Naval Academy (USNA), and Platoon Leaders Class (PLC). Previous studies have also found that officers assessed through E-to-O programs receive higher marks. The difference between USNA and OCC is not statistically significant but all other differences are. Women from "other" commissioning sources score higher than women from all other sources; also, NROTC women were marked higher than USNA women.

In a previous section, we found that—within a given grade—MROs were marked higher by higher ranking reporting seniors (RSs). We hypothesized that MROs assigned to higher ranking RSs may be higher quality on average. However, we find no evidence of this.

Educational credentials

To determine how education enhances officers' performance and whether RSs and ROs see this the same way, we examine the effect of college grade point average (GPA) and major on FRAs and reviewing officer (RO) marks. We conduct this analysis separately among officers who attended a school with a Barron's competitiveness ranking and among officers who accessed through USNA. We observe school

16. Tables 14 and 15 in appendix B contain our regression results.

17. Examples of E-to-O programs are the Marine Enlisted Commissioning Education Program (MECEP), Enlisted Commissioning Program (ECP), and Meritorious Commissioning Program (MCP).

ranking, GPA, and major for approximately one quarter of all officers; however, this smaller sample has the same distribution of race/ethnicity and gender as the full sample.

Among both USNA graduates and other officers, those with higher GPAs tend to get higher FRAs and RO marks.¹⁸ Officers with business-related majors also tend to get higher marks, whereas those with science, engineering, or technology majors tend to get lower marks. Even after controlling for occfield, professional credentials (mostly law degrees) provide the greatest advantage.

Barron's Educational Series publishes a competitiveness scale, categorizing undergraduate programs as "Most Competitive," "Highly Competitive," "Very Competitive," "Competitive," "Less Competitive," "Noncompetitive," or "Special."¹⁹ Alumni of colleges rated as most competitive receive higher FRAs and RO marks than alumni of competitive schools, and alumni of noncompetitive schools receive lower marks. We find it interesting that school quality has a greater effect on RO marks than on the FRA or on the relative value (RV): this suggests a difference between the perspective and priorities of the RS and the RO.

School quality has no direct effect on RV after controlling for RO marks and GPA, suggesting that, although alumni of more competitive schools tend to receive higher FRAs, any advantage that the RS sees is also consistently seen by the RO. The closer observation of the RS reveals nothing additional in the MRO's performance. School quality, however, *does* have additional predictive power for RO marks after controlling for RV and GPA. If an RS is equally satisfied with the performance of two officers, the RO tends to have a higher opinion of the one with a prestigious education. This may suggest that skills or discipline acquired from more competitive schools are more correlated with potential for achievement in broader responsibilities.

18. We include controls for grade and calendar year (see tables 16 through 18 in appendix B).

19. *Barron's Profile of American Colleges 2009*. Hauppauge, NY: Barron's Education Series, Inc., 2008.

Do FitRep marks differ across occupational fields?

In this subsection, we compare the RVs and RO marks received by officers in different occfields with the average for their paygrades and years.²⁰ Because officers in different occfields inherently tend to compete against different peers, we analyze the RV rather than the FRA. For example, logisticians could all tend to receive higher FRAs when evaluated by other logisticians, but this would not raise their RVs unless they also earn higher marks when they are assigned outside their own community. Whether because of differences in quality, level of responsibility, potential career pitfalls, or biases, we find that some occfields are assessed more favorably than others.

Table 5 highlights the most prominent occupations. Occfields with officers receiving significantly higher RVs than average are (in descending order) infantry, aircraft maintenance, armor, logistics, and artillery. Aviators and supply officers tend to receive significantly lower RVs than average.

Table 5. Occfields with highest and lowest RVs and RO marks^a

High RV	High RO	Low RV	Low RO
Infantry	Public affairs	Pilots/NFOs	Pilots/NFOs
Aircraft maintenance	Infantry	Supply	Air control
Armor	Aircraft maintenance		
Logistics	Financial management		

a. Source: CNA analysis of Personnel Evaluation System data and Total Force Data Warehouse data.

Table 5 shows similar results for RO marks. Occfields with officers receiving significantly higher RO marks than average are (in descending order) public affairs, infantry, aircraft maintenance, financial management, aviation logistics, military police, personnel and admin-

20. We first compared each RV or RO mark against what we would predict it to be just from the MRO's paygrade and the calendar year. We computed the difference and averaged these differences for each occfield. For example, on average, public affairs officers receive RO marks 0.21 point higher than we would expect for their paygrade and year.

istration, armor, logistics, communications, intelligence, and artillery. Aviators and air controller officers tend to receive significantly lower RO marks than average. In the next section, we will consider whether differences across occfields may reflect that officers are making evaluations based not only on individual performance but also on general perceptions about other occfields.

Do FitRep marks differ by race and ethnicity?

On average, black and Hispanic officers receive lower FRAs and RO marks than white officers in the same paygrade and year. Specifically, the gap between white and black officers is 0.06 point on the FRA (about one-tenth of the standard deviation of FRAs within a paygrade) and 0.17 point on the RO assessment (about one-fifth of the standard deviation of RO marks within a paygrade). The gap between white and Hispanic officers is smaller—0.03 point on the FRA and 0.10 point on the RO assessment. To put this in perspective, an FRA 0.07 point lower represents a change of one letter on 1 of the 14 PARS; however, these gaps vary considerably depending on whether we control for other characteristics.²¹ First, we control for characteristics we can observe before the officers are commissioned. Then, we control for TBS standing. Finally, we account for occupations and examine whether differences across the Corps are the same as differences within occfields.

Available measures of officer candidate quality before commissioning provide little or no explanation for the difference. Controlling for GCT score explains none of the black-white gap and only 6.1 percent of the Hispanic-white gap in FRA. Controlling for commissioning source actually increases the difference because a higher percentage of black and Hispanic officers are prior enlisted and therefore would be expected to have *higher* marks. Controlling for academic credentials (school competitiveness, college GPA, and college major) explains 12.5 percent of the black-white FRA gap and increases the Hispanic-white gap. Similar results apply to RO marks. Differences in FitRep marks between white, black, and Hispanic officers cannot be

21. Table 19 in appendix B contains our regression results.

attributed to recruiting minority officer candidates with different academic credentials because these credentials explain none of the Hispanic-white gap and only one-eighth of the black-white gap.

TBS performance, however, explains *all* of the difference in FRA (and about two-thirds of the difference in RO marks), for both black and Hispanic officers. In fact, white officers receive slightly *lower* RS marks than do minority officers from the same TBS third. Because TBS performance affects military occupational specialty (MOS) choice, the racial/ethnic gap at TBS has implications for diversity throughout the Corps.

The distribution of occfields is different for black and Hispanic officers than for white officers, and FitRep marks differ by occfield. When we account for this, we find that the black-white gap and the Hispanic-white gap are larger within occfields than across the Marine Corps. On average, black officers receive an FRA 0.12 point lower (about one-fifth of the standard deviation) and an RO mark 0.22 point lower than white officers in the same occfield. The gap between white and Hispanic officers is, once again, smaller—0.07 point on the FRA and 0.13 point on the RO assessment. The gap is larger within occfields because black and Hispanic officers are disproportionately located in such occupations as aircraft maintenance, logistics, and financial management, which otherwise tend to receive high marks. That is, occfield differences would suggest that black and Hispanic officers should receive higher marks on average, rather than lower. Controlling for TBS third explains more than half of the within-occfield gap, for both black and Hispanic officers. Controlling for academic credentials increases the gap.

Summary

As expected, traditional quality measures, such as prior enlisted service, the top third at TBS, or high college GPA, predict higher FitRep marks. Marks tend to differ for officers in different occfields or of different races or ethnicities. The gap between whites and minorities cannot be explained by differences in occfield choice or in academic credentials. In the next section, we look for information to explain the drivers of these differences.

Is there evidence of bias?

Because we do not directly observe an officer's performance, we cannot observe how the FitRep marks evaluating the officer's performance compare with his or her true performance. This makes biased evaluation difficult to prove; however, we *can* follow an officer's FitRep history over time and compare marks from certain evaluators with those from other evaluators. In particular, we observe the demographics and occfields of the reporting senior (RS) and reviewing officer (RO). In this section, we present evidence that we believe is *consistent with* biased evaluation, although there may be other explanations.

Do the race/ethnicity- and gender-match between the RS and MRO affect marks?

FitRep averages (FRAs) differ by the gender of the RS and the Marine reported on (MRO). Female Marine officers receive higher FRAs than male Marine officers—about 0.14 point higher on average (a one-letter change to 2 of the 14 PARS).²² This is consistent with previous studies. Male officers receive higher marks—0.05 point higher on average—from female than from male RSs. Female Marine officers, however, are marked 0.02 point lower by female RSs than by male RSs. This holds when controlling for year, MRO grade, race/ethnicity, marital status, dependents, The Basic School (TBS) third, GCT score, and commissioning source.

FRAs also differ by the race/ethnicity of the RS and the MRO. White RSs give slightly higher marks (0.02 point) to white and Hispanic MROs than to black and “other” race/ethnicity MROs. A previous study found a similar effect among staff sergeants. Conversely, black

22. Tables 20 and 21 in appendix B contain our regression results.

RSs give higher marks (0.06 point) to black MROs than to white MROs; the marks they give to Hispanic and “other” race/ethnicity MROs are not statistically different from those they give to white or black MROs. Hispanic RSs give statistically equivalent marks to Marine officers of all races and ethnicities. Other race/ethnicity RSs grade Hispanics lower than white and other race/ethnicity MROs. Once again, this holds when controlling for year, MRO grade, race/ethnicity, marital status, dependents, TBS third, GCT score, and commissioning source.

Does occfield-match between the RS, RO, and MRO affect marks?

An officer’s actual and perceived performance may depend both on fit with the billet assignment and with the RS’s occupation. Reporting to another officer with the same training may imply less individual responsibility, more mentoring, and/or more accurate observation of performance. We compare the two-digit occfield associated with the MRO’s MOS, the billet MOS, and the RS’s and RO’s MOSs. Some billets, such as B-billets or other specialized assignments, do not match to any particular occfield. To separate these, we labeled a billet as “exceptional” if fewer than 10 percent of the FitReps from that billet apply to MROs whose occfields match the first two digits of the “billet military occupational specialty.” That is, an “exceptional” billet does not appear to be associated with any occfield. We predicted relative values (RVs) and RO marks using paygrade, year, and individual officers’ averages over time. That is, we control for the fact that some officers consistently tend to receive higher marks than others.

It is possible that RSs in one community tend to have a different marking philosophy from those in another community, but are consistent about it, so the RV is unaffected. We want to isolate differences in relative assessment, so we look at the RV rather than the FRA.

Both the RV and the RO mark tend to be lower when the RS and MRO have a matching occfield. On average, the RV is 0.23 point lower, on a scale from 80 to 100. This may be partly driven by greater responsibility when the RS’s occfield does *not* match; the MRO is

likely to be the senior or only officer in his or her occfield in the unit. It may also suggest that RSs are more likely to err on the high side when they have less understanding of the MRO's work. This finding, however, does not hold across all paygrades or across all occfields.

We looked specifically at MROs in the 01, 02, 04, 06, and 44 occfields (personnel, intelligence, logistics, communications, and legal services) because these officers are likely to be assigned to units consisting mostly of Marines in a different occfield. Logistics officers receive RVs 0.26 points higher (on a 20 point scale) when evaluated by other logisticians. There is also marginally significant evidence that first lieutenants (averaged across occfields) and communications officers receive higher RVs when they serve under an RS in the same occfield.²³

RVs and RO marks both have higher variance, and therefore are more informative, when the MRO is in a designated exceptional billet. RO marks also have higher variance when the RO's and MRO's occfields match; that is, the RO gives more informative marks when he or she is more informed about the MRO's duties. RSs appear to do the opposite, however; RVs have lower variance when the RS's and MRO's occfields match.

In summary, there is evidence that some officers benefit from being evaluated by RSs who are less informed about their occupations. This could contribute to higher marks for intelligence officers and judge advocates, because they routinely report to officers outside their occfields. However, logistics officers and communications officers appear to suffer a disadvantage in their RV marks, even while enjoying a similar advantage in RO marks. The effect of the evaluator's experience on FitRep marks suggests that these marks may provide a less efficient signal of an officer's true ability, but that this problem could be mitigated by providing boards with more information.

23. Tables 22 through 26 in appendix B contain our regression results.

Is there evidence of differences in general perceptions of different occfields?

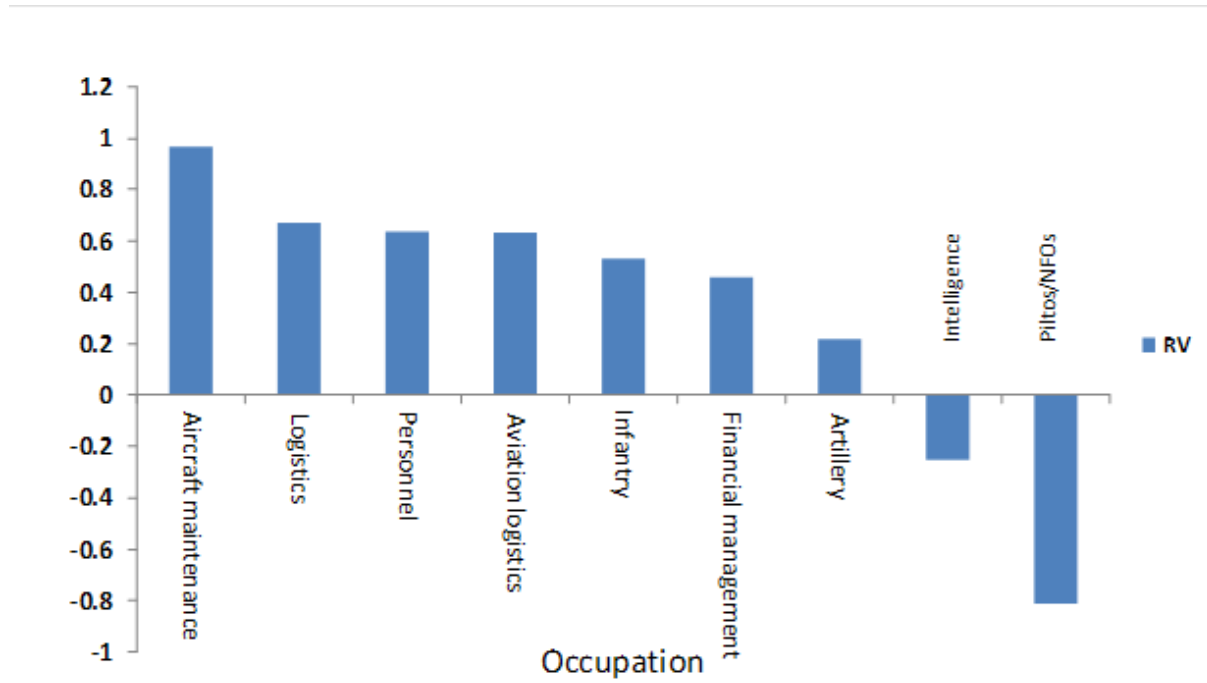
Differences in average marks across occfields could be driven either by bias or by quality differences. Differences in marks from supervisors in the same occfield vice other occfields could be driven by bias or by different responsibilities. To investigate this further, we use TBS third as a proxy for quality, and RS occfield-match as a proxy for responsibilities.

Specifically, we predicted RVs and RO marks using paygrade, year, and TBS third, then looked for prediction errors that differ significantly across occfields. That is, we looked for occfields in which officers consistently tend to receive higher marks than peers from the same TBS third and in the same grade.

Figures 6 and 7 show these results. Officers in aircraft maintenance, logistics, personnel, aviation logistics, infantry, financial management, and artillery receive higher RVs than expected for their TBS third, while aviators and intelligence officers receive lower RVs. The same seven occfields that receive high RVs also receive high RO marks, in addition to public affairs, military police, communications, and supply officers. Only aviators receive lower than average RO marks. The difference from the average RO mark for public affairs officers (0.23) and aviators (-0.2) is comparable to the RO gap between white and black officers (0.17).

The results for aviators are particularly surprising because they typically have their aviation contract in hand before entering TBS, so they know that they have competitive flight training ahead of them. There is anecdotal evidence of aviators “pacing” themselves during TBS as a result. We might therefore expect that they would have lower TBS standings on average, and that their subsequent marks would be high relative to their artificially low TBS third. Instead, we find that they have higher than average TBS standings and lower than average marks.

Figure 6. Difference in average RV from peers in the same paygrade, year of report, and TBS third^a



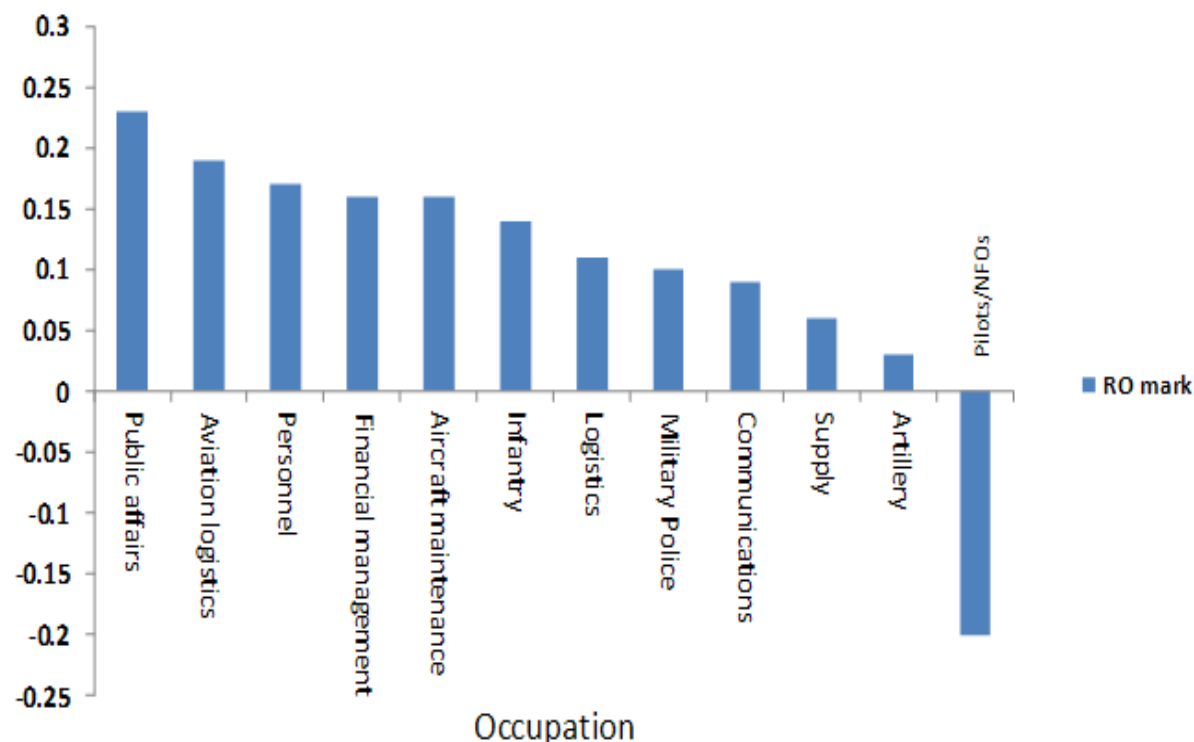
a. Results shown are significant at the 5-percent level (errors clustered by MRO).

Source: CNA analysis of Personnel Evaluation System data and Total Force Data Warehouse data.

We also looked at the effect of having an RO with the same occfield—for MROs of each occfield separately—while restricting attention to MROs with an RS in the same occfield and controlling for the average effect across all occfields of RO-match.²⁴ For example, we tested whether the effect on an infantry officer of having an RO who is also an infantry officer differs significantly from the effect of any officer having an RO in the same occfield, given that the RS is also in the same occfield. In this way, we can identify whether the disadvantage of occfield match (or, conversely, the advantage of an RO from a different occfield) is stronger in some occfields than others and whether, in some occfields, it reverses direction.

24. We also control for paygrade and year. Adding controls for billet match or restricting B-billets from the analysis also had virtually no effect on the results.

Figure 7. Difference in average RO mark from peers in the same paygrade, year of report, and TBS third^a



a. Results shown are significant at the 5-percent level (errors clustered by MRO).

Source: CNA analysis of Personnel Evaluation System data and Total Force Data Warehouse data.

If there is a general negative bias toward an occfield, we would expect that officers in that occfield would receive lower marks than predicted after controlling for TBS third, and that they would receive higher marks from ROs in their same occfield than from others (presumably officers would not be biased against their own occfield). Conversely, if there is a general positive bias toward an occfield, we would expect that officers in that occfield would receive higher marks than predicted, after controlling for TBS third, and that the disadvantage of an RO with a matching occfield would be stronger for these officers than for others.

Four occfields fit this profile. Aviators receive lower RVs and RO marks than expected, and they receive RO marks 0.11 point higher

from fellow aviators than from others.²⁵ We note that unexpectedly low marks for some aviators could be a consequence of missed qualifications, but this does not explain the advantage of having an RO who is also an aviator. Infantry officers and logisticians receive higher RVs and RO marks than expected and have a larger drop in RO marks than usual (0.02 for infantry, 0.12 for logistics) when serving under an RO in the same occfield. There is a similar finding for military police officers: their RVs are not significantly different than expected, but they receive RO marks higher than expected—and 0.25 point lower from ROs who are also military police.

There is some evidence that officers in other occfields could be biased against aviators and toward infantry officers, logisticians, and military police. Aviators receive RVs 0.8 point lower than expected and RO marks 0.2 point lower than expected given their TBS performance, and they are assessed 0.1 point higher by other aviators than by ROs in different occfields. In contrast, logisticians receive RVs 0.7 point higher than expected and RO marks 0.1 point higher than expected given their TBS performance, and they are assessed 0.1 point lower by other logisticians than by ROs in different occfields.

Summary

There is some evidence that RSs and ROs may award marks influenced by biases toward a particular race, gender, or occfield. Specifically, MROs of each gender tend to receive higher FRAs from RSs of the opposite gender, and white RSs tend to award lower FRAs to black MROs and vice versa. Furthermore, RSs and ROs tend to award higher marks to infantry and logistics officers and lower marks to aviators. Further research is needed to understand what drives these apparent biases in FitRep marks and how best to address them.

25. However, aviators receive slightly higher RO marks than expected in the rank of lieutenant colonel.

This page intentionally left blank.

How do subjective comments correlate with FitRep marks?

Competitive boards look not only at relative values (RVs) and reviewing officer (RO) marks but also at subjective comments—recommendations for command, resident schooling, and promotion. FitReps with higher marks should also have strong positive comments.

Because we could not read the section I comments from every FitRep, we focused on the promotion recommendations from a small sample of FitReps. The sample was designed to represent diversity in RV, occfield, and race or ethnicity. We selected 100 FitReps for white officers, 100 for black officers, and 100 for Hispanic officers, all with the rank of captain. Within each set of 100, 20 had an RV less than 84, 20 had an RV between 84 and 88, and so on, up to 20 between 96 and 100. Each set of 20 included 6 reports on ground combat officers, 6 on aviators, and 8 on officers in other occfields (as about half of all officers are in the latter category). All the FitReps in the sample came from deep profiles: each was at least the 10th observed FitRep that the reporting senior (RS) had written on a captain. Given these criteria, the FitReps were chosen by the order of the randomly generated unique identifiers of the Marine reported on (MRO).

We categorized the promotion recommendations into four tiers (without observing the RV or the race or ethnicity of the officer while making these categorizations). Table 6 shows examples of language used in each tier of recommendation. A few FitReps (13 of our original 300) could not be used because the captain had already been selected to major or because the comments continued onto an addendum page that we did not observe.

As table 6 suggests, the precise ordering of any two comments is not always obvious. We do not know whether a particular RS would consider “promote at first opportunity” to be a stronger or weaker state-

ment than “highly recommended for promotion,” or whether all RSs would have the same interpretation of these two comments. We are more confident that “promote ahead of peers” is a stronger endorsement than these two and that “promote with peers” is weaker.

Table 6. Subjective comments classified into tiers of promotion recommendation strength^a

Tier 4	Tier 3	Tier 2	Tier 1
“I do not recommend promotion”	“promote”	“enthusiastically recommended for promotion”	“promote ahead of peers”
“qualified for promotion”	“promote with peers”	“promote at first opportunity”	“groom for highest ranks in Marine Corps”
(nothing)	(implied by recommendation for battalion command)	“highly recommended for promotion”	“my highest recommendation for promotion”
		“promote now”	“a must for promotion”

a. Source: TBS student handout and CNA subjective judgment.

Reports with higher RVs also had stronger promotion recommendations on average. The average RVs for FitReps with a tier 1, 2, 3, or 4 recommendation, respectively, were 94, 92, 88.1, and 87.4.

White officers received stronger promotion recommendations than black or Hispanic officers. Table 7 shows the distribution of recommendations for each race or ethnicity. These differences are statistically significant, controlling for RV. Specifically, black and Hispanic officers each tend to have a recommendation one tier lower than a white officer—with the same RV—28 percent of the time.

Table 7. Strength of promotion recommendation for officers of each race or ethnicity^a

	White captains	Black captains	Hispanic captains
Tier 1	18%	8%	13%
Tier 2	33%	31%	29%
Tier 3	38%	45%	35%
Tier 4	11%	16%	23%
Total number of FitReps	97	93	93

a. Source: CNA analysis of official military personnel files.

This finding is consistent with a possibility that if an RS considers two officers to have similar performance, and therefore awards them similar numerical marks, he or she may have an inadvertent tendency to use stronger language about the white officer.

We also find that ground combat officers receive weaker recommendations than officers in other occfields with the same RV.²⁶ Specifically, ground combat officers tend to have a recommendation one tier lower 25 percent of the time. This suggests that subjective comments for ground officers may not be as high, relative to other occfields, as the numerical marks for those ground officers.

There are at least two reasons to be cautious about this result. The RV is truncated at 80, and some 80s reflect much lower underlying marks (i.e., adverse) than other 80s. When we remove FitReps with an RV of 80 from the analysis, the gap between black or Hispanic and white officers is no longer statistically significant. The other reason is that the analysis excludes comments that continue onto an addendum sheet, and a higher share of these in the sample apply to black or Hispanic officers. Those reports may also have very strong recommendations. Therefore, this relationship between subjective comments and race or ethnicity should continue to be monitored.

Because so much is dependent on the subjective judgment of each RS and RO, we next examine the education these officers receive about FitReps.

26. Table 27 in appendix B contains our regression results.

This page intentionally left blank.

What are Marine officers taught about the FitRep system?

No matter how the PES is structured, its execution will always depend on thousands of individual Marine officers. There is no substitute for their proper understanding of the letter and spirit of the FitRep.

The Basic School (TBS) offers a course on FitReps. It consists of a handout, three lectures, a case study homework assignment, and a discussion group. The first lecture describes the background, purpose, and target audience of the FitRep. It also explains who the Marine reported on (MRO), reporting senior (RS) and reviewing officer (RO) are, how the performance anchored rating scales (PARS) are scored, and how an RS's profile is tracked. The second lecture focuses largely on explaining the importance of the relative value (RV) and on examples of appropriate comments to write about technical skill, leadership, and potential for promotion. The third lecture covers a variety of administrative details of the FitRep form and the submission process.

TBS instruction emphasizes the following points in an attempt to curb inflation and promote the best and most qualified Marines:

- When evaluating your first MRO in a new paygrade, be wary of starting with a high mark.
- Giving everyone the same mark does not help the Marine Corps; you must identify stronger performers.
- Responsible evaluation requires time and care.
- Remember that you are writing for competitive boards, not for the MRO; the FitRep is neither a lever to exert influence nor a counseling tool. Counsel the MRO beforehand, so that the FitRep is not a surprise.

- Consult the Performance Evaluation System manual and your RS profile while writing FitReps.
- Discuss the FitRep with the command and/or senior enlisted Marine before submitting, especially when you are still relatively inexperienced. Do not let senior enlisted influence the level of your marks, only their relative ranking.
- Match comments to marks. Comment specifically on proficiency, leadership, and promotion, and be aware of weak, average, and strong comments in each category.

These recommendations are all helpful and reflect the careful attention TBS pays to teaching this material, but we observed apparent gaps in the instruction. First, neither the handout nor the slides teach Marine officers how the RV is calculated, and both imply that the RV normalizes the marks to fit a “bell curve.” This is not true. The RV rescales the marks but does not change the shape of the distribution except to truncate very low marks. Some officers could have profiles skewed to the left—a cluster of marks that are close together and slightly above the profile average, with a few that are far below—while others could have profiles skewed to the right. The RV does not, and probably should not, adjust for this.

The instructors suggest that an MRO’s marks “should be” increasing over time with the same RS, though there will be exceptions. They do not clearly distinguish between the MRO’s performance truly improving over time, while holding the standard of evaluation constant, and raising the marks with no apparent change in performance. This suggestion could inadvertently lead to inflation.

To our knowledge, this TBS instruction is the only formal education on the topic that active-duty Marines receive. Much of the instruction appears to be quite appropriate and helpful, but officers would probably benefit from receiving refresher training after they have had some experience writing FitReps.

There is broad consensus that junior officers and senior enlisted Marines have little understanding of how promotion boards operate and, therefore, of how FitReps are interpreted and used. In a discussion group at the Expeditionary Warfare School (EWS), captains

expressed frustration that boards place additional weight on specific phrases and that this knowledge is not disseminated. Officer career counselors said that they would like to provide individual instruction about the FitRep system to each student at EWS or the Command and Staff College.

In addition, Evaluation Review (MMER) representatives and former board members believe that the boards should also receive more education on FitReps. Board members must rely on their own experiences to recognize patterns in FitRep marks, such as the fact that detachment officers may receive low marks that are not reflective of their performance when assigned to a different squadron than the one with which they trained. They may also have little knowledge of performance evaluation review boards (PERBs), which can change or delete portions of a Marine's record.

Manpower Management Division is currently engaged in developing training to help Marines understand how boards interpret and use FitReps. However, proper implementation of FitRep training will require cooperation from Training and Education Command.

Because the ultimate purpose of the FitRep system is to inform boards, our concluding analysis in the next section is of the way FitRep information is displayed to board members.

This page intentionally left blank.

How might boards' view of FitRep results contribute to the boards' confusion?

Board members have electronic access to all FitReps that have not been removed from an officer's record. The display system through which they access this information includes a "briefing guide" that tabulates relative values (RVs) and reviewing officer (RO) marks. This screen shows the number of FitReps an officer has received that fell in the upper third ($RV > 93.33$), middle third ($86.66 < RV < 93.34$), and lower third ($RV < 86.67$) of the reporting senior's (RS's) profile. It also shows the number of RO assessments—from this officer's ROs to other Marines reported on (MROs) in the same grade—that were above, with, and below the mark they gave this officer. For both RVs and RO marks, the results are tabulated two ways: at the time each report was processed ("At") and now ("Cum."). Table 8 illustrates this tabulation.

Table 8. View of RV and RO mark tabulation in briefing guide for a fictional officer

	At	Cum.
RV Summary		
Upper third	2	2
Middle third	3	5
Lower third	2	3
N/A	4	1
RO Assessment		
Above	10	24
With	19	35
Below	16	22

For example, suppose Major Green is being considered for promotion to lieutenant colonel. One of the FitReps he received as a captain

had an RV of 95 at the time it was processed. The RO assessment on this FitRep was a 5, and that RO had at that time awarded two 4s and two 5s to captains. In the “At” column, this FitRep is counted as an upper third (for the RV of 95)—the number in the “Upper third” row increases by one—and contributes one to the “With” row (for the other 5 given to another captain) and two to the “Below” row (for the two 4s given to other captains). The other numbers in the “At” column (e.g., 2 instead of 1 in the “Upper third” row, non-zero values in the other RV summary rows) reflect other FitReps that Major Green has received.

Since then, the RS and RO have both tended to give higher marks to captains. This FitRep is now in the middle third of the RS’s profile, and the RO’s complete history includes one 3, four 4s, seven 5s, and two 6s. In the “Cum.” column, this FitRep adds one to the “Middle third” row (for an RV between 86 and 93) and adds two to “Above” (for the two 6s), six to “With” (for the six other 5s), and five to “Below” (for the one 3 and four 4s).

Analysts at the Marine Corps’ Operations Analysis Division (OAD) have recognized some potentially confusing differences between how RVs and RO marks are tabulated. Larger numbers in the top row of the RV summary are good, whereas larger numbers in the bottom row of the RO assessment are good. Each FitRep from the MRO’s file receives a weight of one in the RV summary—regardless of the time period covered, and as long as it generated an RV—but the RO assessment weights them according to how many other FitReps a particular RO has written for a grade. If one RO has assessed hundreds of other officers in the same paygrade, then, for each of those officers, the assessment he or she received from this RO could dwarf all other FitReps received over the course of his or her career.

The OAD analysts compare the RO assessment, as it would appear in a briefing guide, with the average of RVs over an officer’s career. Because the RV is scaled between 80 and 100, they make the comparison by equating each “Above” to 80, each “With” to 90, and each “Below” to 100. They then average these numbers and subtract from the RV average. The difference between the converted RO history and the RV history has a skewed distribution; it is more common for

the RV history to be much more positive than the RO history. There are officers who have tended to consistently receive high RVs, even while their ROs have tended to rate other officers above them. There do not appear to be cases tilted strongly the other way. This suggests that, in some cases, a low RO mark significantly alters an officer’s history as displayed to the board, in a way that is not consistent with other information about the officer’s career.

Table 9 illustrates a more intuitive method for tabulating RO marks. Specifically, for each FitRep that the officer has received with an observed RO mark, the screen could display whether that mark was above, equal to, or below the RO’s median for that paygrade. Now, both the RV and RO tabulations are structured in such a way that higher numbers at the top are better and that each report is counted once.

Table 9. Proposed alternative view of RV and RO mark tabulation in briefing guide

	At	Cum.
RV Summary		
Upper third	2	2
Middle third	3	5
Lower third	2	3
N/A	4	1
RO Assessment		
Above median	3	3
With median	5	6
Below median	3	2

Neither the briefing guide nor the individual FitReps record how closely the RO observed the MRO’s performance. The RO has the option to declare insufficient observation and not enter a comparative assessment. However, in a discussion group at the Command and Staff College, majors expressed a belief that often an RO makes a comparative assessment based on secondhand information and very little direct observation. There is no option to declare that this mark is based on indirect observation.

The FitRep and the master brief sheet both report the time period that each report covers. Officer career counselors observed that the RO may have changed during that period, yet there is no place to record this. A board may place a lot of weight on the RO mark from a report covering 13 months, not knowing that the RO observed only one month. Majors at Command and Staff College also noted this problem.

Before concluding, we summarize subject matter expert input regarding the process of completing and submitting FitReps.

How might the processes for completing and submitting FitReps be improved?

Now that FitReps are entered and submitted through an automated system (A-PES), there may be ways to gain further advantages from this system. For example, the PES manual states that reviewing officers (ROs) should not submit a mark lower than what they gave the same Marine previously, unless there has been a change in performance. Career counselors and representatives from Manpower Management Evaluation Review (MMER) suggest that drops in RO marks are usually unintended. Theoretically, an automated notification could appear when an RO is about to submit a lower mark than before, thereby reducing errors.

Majors at Command and Staff College believe that the system should require more MRO involvement. When the approximate end date of an evaluation period is known in advance, there may be a way for A-PES to prompt the reporting senior (RS) to conduct a FitRep-centered discussion with the Marine reported on (MRO) midway through the evaluation period. Another automated prompt could remind the MRO to write a billet description and list of billet accomplishments 30 days before the end of the period. We recognize that not all units may have access to these prompts when they are sent.

As the RS's immediate superior, the RO is well positioned to counsel the RS on his or her responsibilities. Captains at EWS and majors at Command and Staff College feel it is helpful when the RO monitors the RS's profile and discusses with the RS how a Marine's performance compares with the performance of others in the RS's profile. A-PES could help ROs to monitor the RSs under them by alerting them as to which MROs have FitReps due.

Section H, a performance anchored rating scale for fulfillment of evaluation responsibilities, is a potential mechanism for accountabil-

ity. Majors believe that officers who unfairly harm other Marines' careers through negligence as RSs and ROs should be penalized. Training could emphasize the importance of this section.

Adverse FitReps must be signed by a third sighter who is senior to the RO. Currently, most third sighters provide only the required statement certifying that the FitRep is procedurally correct. MMER representatives note that the third sighter is in a better position than the review board to look into the facts of the case and to help adjudicate it. Career counselors believe that the third sighter could be especially helpful when the RS and/or RO are not Marine officers.

RSs can look up their profiles, containing all FitRep averages (FRAs) they have given and the associated relative values (RVs) into which these were converted. However, they do not see the RV for a report until after they have submitted it. Some stakeholders wonder whether A-PES could calculate and display the RV *before* the RS submits the report. Representatives from MMER and from MMOA, as well as captains and majors, believe that viewing the RV before submission would reduce errors. Displaying the RV on the report would also ensure that the MRO knows his or her RV.

However, calculating and updating the RV for each report as it is being written would be problematic in at least two ways. First, representatives from MMSB assure us that this would significantly strain the capacity of the A-PES computing infrastructure. Second, if the RS is submitting reports for more than one officer at the same time, or within days of each other, the RV calculated as the FitRep is submitted would immediately become outdated.

Conclusions

Our analyses have revealed multiple strengths of the current FitRep system and potential issues to be monitored or addressed. We summarize here the evidence of the system's successes, our recommendations for further training and clearer presentation of FitRep marks, and issues that we believe require further monitoring and study.

Successes of the FitRep system

There is *no* evidence that grade inflation among officers is rampant at an aggregate level across the Marine Corps. FitRep averages (FRAs) rose slightly through FY 2003 but have since declined.

The inclusion of both reporting senior (RS) and reviewing officer (RO) marks appears to be helpful. Both provide valuable information because each helps to predict the other.

FitRep marks are consistent with other indicators of officer quality. Marine officers commissioned through E-to-O programs, with higher college GPAs, or finishing in the top third of their TBS class tend to receive higher marks. In addition, officers who were rated lower in the final years of the previous FitRep system continued to be rated lower under the new system.

Subject matter experts agree that the current system usually results in promotion of the best and most qualified officers.

Additional training for RSs, ROs, and boards

To our knowledge, the only official instruction that officers receive about FitReps occurs at TBS, before they have any experience writing FitReps. They are told that boards are the target audience, but they have little insight into a board's perspective. We recommend that professional military education (PME) for junior officers and senior

enlisted Marines include instruction about the board process, including a view of boardroom tools, examples of boards discussing fictional officers, and subjective comments that boards specifically want to see.

RSs tend to award higher marks the longer they have been evaluating a paygrade and, particularly, the longer they have been evaluating a Marine reported on (MRO). Current training may contribute to this. We recommend that TBS and refresher training emphasize that each evaluating period should be evaluated independently.

RO marks are not consistent with the suggested distribution and have a higher average at higher paygrades. TBS instruction does not include any information about how to properly fill out the RO assessment. We recommend that the Marine Corps provide training for ROs and emphasize that RO marks are compiled separately for each paygrade and should reflect a relative assessment for that grade.

Current instruction on FitReps does not teach officers how the relative value (RV) is calculated, and it incorrectly suggests that the RV converts FRAs to a normal distribution (the bell curve). We recommend that TBS and refresher training teach officers the true interpretation of the RV.

As RSs begin writing FitReps on a new paygrade, they follow a specific pattern in establishing their profiles. These early marks differ from those in a mature profile. We recommend education for board members encouraging them to view FitReps from early in an RS's profile with some caution. We also recommend that the Marine Corps consider increasing the minimum number of FitReps—or of unique MROs—in an RS profile required to generate an RV.

There is evidence of room-to-grow and welcome-to-the-grade marks masking an officer's true performance. We recommend that boards view the first FitRep an officer receives in a new paygrade or from a new RS with caution and that training explicitly discourage welcome-to-the-grade marks.

Board members vary in knowledge and experience. We recommend that the Marine Corps capture lessons learned from officers with significant board experience.

Clearer and more informative presentation of RO marks

The briefing guide for board members tabulates RO marks in a way that is potentially confusing and weighted inefficiently. We recommend that it tabulate RO marks in a manner more similar to its tabulation of RVs.

Some ROs have direct observation of the MRO, while others make an assessment using indirect information. We recommend an alternative box for the RO to check, indicating indirect observation.

The FitRep and master brief sheet show the span of time the RS observed the MRO, which may not be the same as the span of observation for the RO. We recommend that the master brief sheet note the period of RO observation.

Issues for further monitoring and study

FRAs are becoming somewhat less varied and potentially less informative over time. This is a worrisome trend that the Marine Corps should monitor to determine whether it is continuing.

Observable characteristics of officer candidates prior to commissioning explain little to none of the difference in FitRep marks between white, black, and Hispanic officers, whereas TBS performance differs significantly by race and is a key predictor of the subsequent FitRep gap. We suggest further research into the gap in TBS performance between white, black, and Hispanic officers.

There is evidence that white RSs tend to award *slightly* lower FRAs to black MROs and vice versa—while controlling for other observable characteristics—and that male and female officers each tend to receive higher marks from RSs of the opposite gender. Further study should monitor this pattern and seek to identify underlying causes.

Aviators appear to receive lower marks than other officers of the same quality, and lower marks from ROs in occfields other than their own. The opposite appears to be true for infantry, logistics, and military police officers. We suggest further study to identify underlying causes.

From a limited sample, white officers appear to receive stronger recommendations for promotion than black or Hispanic officers with the same RV. We recommend continued monitoring of the relationship between subjective comments and race or ethnicity.

School quality appears to have a less robust effect than college GPA on performance. Officer selection officers and commissioning boards may consider placing more emphasis on GPA than on school quality.

Appendix A: Blank FitRep Form

Pages 66 through 71 display the five-page FitRep form and a blank addendum page.

COMMANDANT'S GUIDANCE

DO NOT STAPLE
THIS FORM

The completed fitness report is the most important information component in manpower management. It is the primary means of evaluating a Marine's performance and is the Commandant's primary tool for the selection of personnel for promotion, augmentation, resident schooling, command, and duty assignments. Therefore, the completion of this report is one of an officer's most critical responsibilities. Inherent in this duty is the commitment of each Reporting Senior and Reviewing Officer to ensure the integrity of the system by giving close attention to accurate marking and timely reporting. Every officer serves a role in the scrupulous maintenance of this evaluation system, ultimately important to both the individual and the Marine Corps. Inflationary markings only serve to dilute the actual value of each report. Reviewing Officers will not concur with inflated reports.

A. ADMINISTRATIVE INFORMATION

1. Marine Reported On:

a. Last Name	b. First Name	c. MI	d. SSN	e. Grade	f. DOR	g. PMOS	h. BILMOS

2. Organization:

a. MCC	b. RUC	c. Unit Description

3. Occasion and Period Covered:

a. OCC	b. From	To	c. Type

4. Duty Assignment (descriptive title):

5. Special Case:

a. Adverse	b. Not Observed	c. Extended
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

6. Marine Subject Of:

a. Commendatory Material	b. Derogatory Material	c. Disciplinary Action
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

7. Recommended For Promotion:

a. Yes	b. No	c. N/A
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

8. Special Information:

a. QUAL	d. HT(in.)	g. Reserve Component
<input type="text"/>	<input type="text"/>	<input type="text"/>
b. PFT	e. WT	h. Future Use
<input type="text"/>	<input type="text"/>	<input type="text"/>
c. Status	f. Body Fat	i. Future Use
<input type="text"/>	<input type="text"/>	<input type="text"/>

9. Duty Preference:

a. Code	b. Descriptive Title

1st	
2nd	
3rd	

10. Reporting Senior:

a. Last Name	b. Init	c. Service	d. SSN	e. Grade	f. Duty Assignment

11. Reviewing Officer:

a. Last Name	b. Init	c. Service	d. SSN	e. Grade	f. Duty Assignment

B. BILLET DESCRIPTION

C. BILLET ACCOMPLISHMENTS

1. Marine Reported On:				2. Occasion and Period Covered:			
a. Last Name		b. First Name		c. MI	d. SSN	a. OCC	b. From To
D. MISSION ACCOMPLISHMENT 1. PERFORMANCE. Results achieved during the reporting period. How well those duties inherent to a Marine's billet, plus all additional duties, formally and informally assigned, were carried out. Reflects a Marine's aptitude, competence, and commitment to the unit's success above personal reward. Indicators are time and resource management, task prioritization, and tenacity to achieve positive ends consistently.							
ADV	Meets requirements of billet and additional duties. Aptitude, commitment, and competence meet expectations. Results maintain status quo.	Consistently produces quality results while measurably improving unit performance. Habitually makes effective use of time and resources; improves billet procedures and products. Positive impact extends beyond billet expectations.	Results far surpass expectations. Recognizes and exploits new resources; creates opportunities. Emulated; sought after as an expert with influence beyond unit. Impact significant; innovative approaches to problems produce significant gains in quality and efficiency.				N/O
A	B	C	D	E	F	G	H
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2. PROFICIENCY. Demonstrates technical knowledge and practical skill in the execution of the Marine's overall duties. Combines training, education and experience. Translates skills into actions which contribute to accomplishing tasks and missions. Imparts knowledge to others. Grade dependent.							
ADV	Competent. Possesses the requisite range of skills and knowledge commensurate with grade and experience. Understands and articulates basic functions related to mission accomplishment.	Demonstrates mastery of all required skills. Expertise, education and experience consistently enhance mission accomplishment. Innovative troubleshooter and problem solver. Effectively imparts skills to subordinates.	True expert in field. Knowledge and skills impact far beyond those of peers. Translates broad-based education and experience into forward thinking, innovative actions. Makes immeasurable impact on mission accomplishment. Peerless teacher, selflessly imparts expertise to subordinates, peers, and seniors.				N/O
A	B	C	D	E	F	G	H
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
JUSTIFICATION:							
E. INDIVIDUAL CHARACTER 1. COURAGE. Moral or physical strength to overcome danger, fear, difficulty or anxiety. Personal acceptance of responsibility and accountability, placing conscience over competing interests regardless of consequences. Conscious, overriding decision to risk bodily harm or death to accomplish the mission or save others. The will to persevere despite uncertainty.							
ADV	Demonstrates inner strength and acceptance of responsibility commensurate with scope of duties and experience. Willing to face moral or physical challenges in pursuit of mission accomplishment.	Guided by conscience in all actions. Proven ability to overcome danger, fear, difficulty or anxiety. Exhibits bravery in the face of adversity and uncertainty. Not deterred by morally difficult situations or hazardous responsibilities.	Uncommon bravery and capacity to overcome obstacles and inspire others in the face of moral dilemma or life-threatening danger. Demonstrated under the most adverse conditions. Selfless. Always places conscience over competing interests regardless of physical or personal consequences.				N/O
A	B	C	D	E	F	G	H
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2. EFFECTIVENESS UNDER STRESS. Thinking, functioning and leading effectively under conditions of physical and/or mental pressure. Maintaining composure appropriate for the situation, while displaying steady purpose of action, enabling one to inspire others while continuing to lead under adverse conditions. Physical and emotional strength, resilience and endurance are elements.							
ADV	Exhibits discipline and stability under pressure. Judgment and effective problem-solving skills are evident.	Consistently demonstrates maturity, mental agility and willpower during periods of adversity. Provides order to chaos through the application of intuition, problem-solving skills, and leadership. Composure reassures others.	Demonstrates seldom-matched presence of mind under the most demanding circumstances. Stabilizes any situation through the resolute and timely application of direction, focus and personal presence.				N/O
A	B	C	D	E	F	G	H
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3. INITIATIVE. Action in the absence of specific direction. Seeing what needs to be done and acting without prompting. The instinct to begin a task and follow through energetically on one's own accord. Being creative, proactive and decisive. Transforming opportunity into action.							
ADV	Demonstrates willingness to take action in the absence of specific direction. Acts commensurate with grade, training and experience.	Self-motivated and action-oriented. Foresight and energy consistently transform opportunity into action. Develops and pursues creative, innovative solutions. Acts without prompting. Self-starter.	Highly motivated and proactive. Displays exceptional awareness of surroundings and environment. Uncanny ability to anticipate mission requirements and quickly formulate original, far-reaching solutions. Always takes decisive, effective action.				N/O
A	B	C	D	E	F	G	H
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
JUSTIFICATION:							
NAVMC 10835B (Rev. 4-03) (P)							

1. Marine Reported On:				2. Occasion and Period Covered:			
a. Last Name		b. First Name		c. MI	d. SSN	a. OCC	b. From To

F. LEADERSHIP							
1. LEADING SUBORDINATES. The inseparable relationship between leader and led. The application of leadership principles to provide direction and motivate subordinates. Using authority, persuasion and personality to influence subordinates to accomplish assigned tasks. Sustaining motivation and morale while maximizing subordinates' performance.							
ADV	Engaged; provides instructions and directs execution. Seeks to accomplish mission in ways that sustain motivation and morale. Actions contribute to unit effectiveness.	Achieves a highly effective balance between direction and delegation. Effectively tasks subordinates and clearly delineates standards expected. Enhances performance through constructive supervision. Fosters motivation and enhances morale. Builds and sustains teams that successfully meet mission requirements. Encourages initiative and candor among subordinates.	Promotes creativity and energy among subordinates by striking the ideal balance of direction and delegation. Achieves highest levels of performance from subordinates by encouraging individual initiative. Engenders willing subordination, loyalty, and trust that allow subordinates to overcome their perceived limitations. Personal leadership fosters highest levels of motivation and morale, ensuring mission accomplishment even in the most difficult circumstances.	N/O			
A	B	C	D	E	F	G	H
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2. DEVELOPING SUBORDINATES. Commitment to train, educate, and challenge all Marines regardless of race, religion, ethnic background, or gender. Mentorship. Cultivating professional and personal development of subordinates. Developing team players and esprit de corps. Ability to combine teaching and coaching. Creating an atmosphere tolerant of mistakes in the course of learning.							
ADV	Maintains an environment that allows personal and professional development. Ensures subordinates participate in all mandated development programs.	Develops and institutes innovative programs, to include PME, that emphasize personal and professional development of subordinates. Challenges subordinates to exceed their perceived potential thereby enhancing unit morale and effectiveness. Creates an environment where all Marines are confident to learn through trial and error. As a mentor, prepares subordinates for increased responsibilities and duties.	Widely recognized and emulated as a teacher, coach and leader. Any Marine would desire to serve with this Marine because they know they will grow personally and professionally. Subordinate and unit performance far surpassed expected results due to MRO's mentorship and team building talents. Attitude toward subordinate development is infectious, extending beyond the unit.	N/O			
A	B	C	D	E	F	G	H
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3. SETTING THE EXAMPLE. The most visible facet of leadership: how well a Marine serves as a role model for all others. Personal action demonstrates the highest standards of conduct, ethical behavior, fitness, and appearance. Bearing, demeanor, and self-discipline are elements.							
ADV	Maintains Marine Corps standards for appearance, weight, and uniform wear. Sustains required level of physical fitness. Adheres to the tenets of the Marine Corps core values.	Personal conduct on and off duty reflects highest Marine Corps standards of integrity, bearing and appearance. Character is exceptional. Actively seeks self-improvement in wide-ranging areas. Dedication to duty and professional example encourage others' self-improvement efforts.	Model Marine, frequently emulated. Exemplary conduct, behavior, and actions are tone-setting. An inspiration to subordinates, peers, and seniors. Remarkable dedication to improving self and others.	N/O			
A	B	C	D	E	F	G	H
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. ENSURING WELL-BEING OF SUBORDINATES. Genuine interest in the well-being of Marines. Efforts enhance subordinates' ability to concentrate focus on unit mission accomplishment. Concern for family readiness is inherent. The importance placed on welfare of subordinates is based on the belief that Marines take care of their own.							
ADV	Deals confidently with issues pertinent to subordinate welfare and recognizes suitable courses of action that support subordinates' well-being. Applies available resources, allowing subordinates to effectively concentrate on the mission.	Instills and/or reinforces a sense of responsibility among junior Marines for themselves and their subordinates. Actively fosters the development of and uses support systems for subordinates which improve their ability to contribute to unit mission accomplishment. Efforts to enhance subordinate welfare improve the unit's ability to accomplish its mission.	Noticeably enhances subordinates well-being, resulting in a measurable increase in unit effectiveness. Maximizes unit and base resources to provide subordinates with the best support available. Proactive approach serves to energize unit members to "take care of their own," thereby correcting potential problems before they can hinder subordinates' effectiveness. Widely recognized for techniques and policies that produce results and build morale. Builds strong family atmosphere. Puts motto "Mission first, Marines always," into action.	N/O			
A	B	C	D	E	F	G	H
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5. COMMUNICATION SKILLS. The efficient transmission and receipt of thoughts and ideas that enable and enhance leadership. Equal importance given to listening, speaking, writing, and critical reading skills. Interactive, allowing one to perceive problems and situations, provide concise guidance, and express complex ideas in a form easily understood by everyone. Allows subordinates to ask questions, raise issues and concerns and venture opinions. Contributes to a leader's ability to motivate as well as counsel.							
ADV	Skilled in receiving and conveying information. Communicates effectively in performance of duties.	Clearly articulates thoughts and ideas, verbally and in writing. Communication in all forms is accurate, intelligent, concise, and timely. Communicates with clarity and verve, ensuring understanding of intent or purpose. Encourages and considers the contributions of others.	Highly developed facility in verbal communication. Adept in composing written documents of the highest quality. Combines presence and verbal skills which engender confidence and achieve understanding irrespective of the setting, situation, or size of the group addressed. Displays an intuitive sense of when and how to listen.	N/O			
A	B	C	D	E	F	G	H
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
JUSTIFICATION:							

NAVMC 10835C (Rev.4-03) (P A-PES 5.1.1.0)	PAGE 3 OF 5
---	-------------

1. Marine Reported On:				2. Occasion and Period Covered:				
a. Last Name		b. First Name		c. MI		d. SSN		
a. OCC		b. From		c. To				
G. INTELLECT AND WISDOM								
1. PROFESSIONAL MILITARY EDUCATION (PME). Commitment to intellectual growth in ways beneficial to the Marine Corps. Increases the breadth and depth of warfighting and leadership aptitude. Resources include resident schools; professional qualifications and certification processes; nonresident and other extension courses; civilian educational institution coursework; a personal reading program that includes (but is not limited to) selections from the Commandant's Reading List; participation in discussion groups and military societies; and involvement in learning through new technologies.								
ADV	Maintains currency in required military skills and related developments. Has completed or is enrolled in appropriate level of PME for grade and level of experience. Recognizes and understands new and creative approaches to service issues. Remains abreast of contemporary concepts and issues.	PME outlook extends beyond MOS and required education. Develops and follows a comprehensive personal program which includes broadened professional reading and/or academic course work; advances new concepts and ideas.	Dedicated to life-long learning. As a result of active and continuous efforts, widely recognized as an intellectual leader in professionally related topics. Makes time for study and takes advantage of all resources and programs. Introduces new and creative approaches to services issues. Engages in a broad spectrum of forums and dialogues.					N/O
A	B	C	D	E	F	G	H	
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
2. DECISION MAKING ABILITY. Viable and timely problem solution. Contributing elements are judgment and decisiveness. Decisions reflect the balance between an optimal solution and a satisfactory, workable solution that generates tempo. Decisions are made within the context of the commander's established intent and the goal of mission accomplishment. Anticipation, mental agility, intuition, and success are inherent.								
ADV	Makes sound decisions leading to mission accomplishment. Actively collects and evaluates information and weighs alternatives to achieve timely results. Confidently approaches problems; accepts responsibility for outcomes.	Demonstrates mental agility; effectively prioritizes and solves multiple complex problems. Analytical abilities enhanced by experience, education, and intuition. Anticipates problems and implements viable, long-term solutions. Steadfast, willing to make difficult decisions.	Widely recognized and sought after to resolve the most critical, complex problems. Seldom matched analytical and intuitive abilities; accurately foresees unexpected problems and arrives at well-timed decisions despite fog and friction. Completely confident approach to all problems. Masterfully strikes a balance between the desire for perfect knowledge and greater tempo.					N/O
A	B	C	D	E	F	G	H	
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
3. JUDGMENT. The discretionary aspect of decision making. Draws on core values, knowledge, and personal experience to make wise choices. Comprehends the consequences of contemplated courses of action.								
ADV	Majority of judgments are measured, circumspect, relevant and correct.	Decisions are consistent and uniformly correct, tempered by consideration of their consequences. Able to identify, isolate and assess relevant factors in the decision making process. Opinions sought by others. Subordinates personal interest in favor of impartiality.	Decisions reflect exceptional insight and wisdom beyond this Marine's experience. Counsel sought by all; often an arbiter. Consistent, superior judgment inspires the confidence of seniors.					N/O
A	B	C	D	E	F	G	H	
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
JUSTIFICATION:								
H. FULFILLMENT OF EVALUATION RESPONSIBILITIES								
1. EVALUATIONS. The extent to which this officer serving as a reporting official conducted, or required others to conduct, accurate, uninflated, and timely evaluations.								
ADV	Occasionally submitted untimely or administratively incorrect evaluations. As RS, submitted one or more reports that contained inflated markings. As RO, concurred with one or more reports from subordinates that were returned by HQMC for inflated marking.	Prepared uninflated evaluations which were consistently submitted on time. Evaluations accurately described performance and character. Evaluations contained no inflated markings. No reports returned by RO or HQMC for inflated marking. No subordinates' reports returned by HQMC for inflated marking. Few, if any, reports were returned by RO or HQMC for administrative errors. Section Cs were void of superlatives. Justifications were specific, verifiable, substantive, and where possible, quantifiable and supported the markings given.	No reports submitted late. No reports returned by either RO or HQMC for administrative correction or inflated markings. No subordinates' reports returned by HQMC for administrative correction or inflated markings. Returned procedurally or administratively incorrect reports to subordinates for correction. As RO nonconcurred with all inflated reports.					N/O
A	B	C	D	E	F	G	H	
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
JUSTIFICATION:								

1. Marine Reported On:				2. Occasion and Period Covered:			
a. Last Name	b. First Name	c. MI	d. SSN	a. OCC	b. From	To	
I. DIRECTED AND ADDITIONAL COMMENTS							
J. CERTIFICATION							
1. I CERTIFY that to the best of my knowledge and belief all entries made hereon are true and without prejudice or partiality and that I have provided a signed copy of this report to the Marine Reported on.				<div style="display: flex; justify-content: space-between;"> <div style="border-bottom: 1px solid black; width: 150px;"></div> <div>(Signature of Reporting Senior)</div> </div>			
				<div style="display: flex; justify-content: space-between;"> <div style="border-bottom: 1px solid black; width: 150px;"></div> <div>(Date in YYYYMMDD format)</div> </div>			
2. I ACKNOWLEDGE the adverse nature of this report and							
<input type="checkbox"/> I have no statement to make <input type="checkbox"/> I have attached a statement				<div style="display: flex; justify-content: space-between;"> <div style="border-bottom: 1px solid black; width: 150px;"></div> <div>(Signature of Marine Reported On)</div> </div>			
				<div style="display: flex; justify-content: space-between;"> <div style="border-bottom: 1px solid black; width: 150px;"></div> <div>(Date in YYYYMMDD format)</div> </div>			
K. REVIEWING OFFICER COMMENTS							
1. OBSERVATION: <input type="checkbox"/> Sufficient <input type="checkbox"/> Insufficient				2. EVALUATION: <input type="checkbox"/> Concur <input type="checkbox"/> Do Not Concur			
3. COMPARATIVE ASSESSMENT: Provide a comparative assessment of potential by placing an "X" in the appropriate box. In marking the comparison, consider all Marines of this grade whose professional abilities are known to you personally.	DESCRIPTION		COMPARATIVE ASSESSMENT				
	THE EMINENTLY QUALIFIED MARINE						
	ONE OF THE FEW						
	EXCEPTIONALLY QUALIFIED MARINES						
	ONE OF THE MANY HIGHLY QUALIFIED						
	PROFESSIONALS WHO FORM THE MAJORITY OF THIS GRADE						
	A QUALIFIED MARINE						
UNSATISFACTORY							
4. REVIEWING OFFICER COMMENTS: Amplify your comparative assessment mark; evaluate potential for continued professional development to include: promotion, command, assignment, resident PME, and retention; and put Reporting Senior marks and comments in perspective.							
5. I CERTIFY that to the best of my knowledge and belief all entries made hereon are true and without prejudice or partiality.				<div style="display: flex; justify-content: space-between;"> <div style="border-bottom: 1px solid black; width: 150px;"></div> <div>(Signature of Reviewing Officer)</div> </div>			
				<div style="display: flex; justify-content: space-between;"> <div style="border-bottom: 1px solid black; width: 150px;"></div> <div>(Date in YYYYMMDD format)</div> </div>			
6. I ACKNOWLEDGE the adverse nature of this report and							
<input type="checkbox"/> I have no statement to make <input type="checkbox"/> I have attached a statement				<div style="display: flex; justify-content: space-between;"> <div style="border-bottom: 1px solid black; width: 150px;"></div> <div>(Signature of Marine Reported On)</div> </div>			
				<div style="display: flex; justify-content: space-between;"> <div style="border-bottom: 1px solid black; width: 150px;"></div> <div>(Date in YYYYMMDD format)</div> </div>			
L. ADDENDUM PAGE							
ADDENDUM PAGE ATTACHED: <input type="checkbox"/> YES							
<div style="display: flex; justify-content: space-between;"> <div>NAVMC 10835E (Rev. 4-03) (P A-PES 5.1.1.0)</div> <div>PAGE 5 OF 5</div> </div>							

USMC FITNESS REPORT NAVMC 11297 (Rev. 4-03) (P)						DO NOT STAPLE THIS FORM	
ADDENDUM PAGE							
A. PURPOSE							
1. Marine Reported On:					2. Occasion and Period Covered:		
a. Last Name	b. First Name	c. M.I.	d. SSN	e. Grade	a. OCC	b. From	To
3. Purpose:							
a. Continuation of Comments Justification Section I RO	b. Accelerated Promotion Justification	c. Adverse Report MRO Statement 3rd Officer Sighter		d. Admin Review	e. Supplemental Material	f. HQMC Use	
<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/>		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
B. TEXT							
C. SUBMITTED BY							
1. a. Last Name	b. First Name	c. MI	2. SSN	3. Service	4. Grade		
_____ Signature				<div style="display: flex; align-items: center;"> <div style="border: 1px solid black; width: 20px; height: 20px; margin-right: 5px;"></div> <div style="border: 1px solid black; width: 20px; height: 20px; margin-right: 5px;"></div> <div style="border: 1px solid black; width: 20px; height: 20px; margin-right: 5px;"></div> <div style="border: 1px solid black; width: 20px; height: 20px; margin-right: 5px;"></div> <div style="border: 1px solid black; width: 20px; height: 20px; margin-right: 5px;"></div> <div style="border: 1px solid black; width: 20px; height: 20px; margin-right: 5px;"></div> <div style="border: 1px solid black; width: 20px; height: 20px; margin-right: 5px;"></div> <div style="border: 1px solid black; width: 20px; height: 20px;"></div> </div> (Date in YYYYMMDD format)			
D. GENERAL/SENIOR OFFICER ADVERSE REPORT SIGHTING							
1. a. Last Name	b. First Name	c. MI	2. SSN	3. Service	4. Grade		
5. Title							
_____ Signature				<div style="display: flex; align-items: center;"> <div style="border: 1px solid black; width: 20px; height: 20px; margin-right: 5px;"></div> <div style="border: 1px solid black; width: 20px; height: 20px; margin-right: 5px;"></div> <div style="border: 1px solid black; width: 20px; height: 20px; margin-right: 5px;"></div> <div style="border: 1px solid black; width: 20px; height: 20px; margin-right: 5px;"></div> <div style="border: 1px solid black; width: 20px; height: 20px; margin-right: 5px;"></div> <div style="border: 1px solid black; width: 20px; height: 20px; margin-right: 5px;"></div> <div style="border: 1px solid black; width: 20px; height: 20px; margin-right: 5px;"></div> <div style="border: 1px solid black; width: 20px; height: 20px;"></div> </div> (Date in YYYYMMDD format)			
PAGE <input type="checkbox"/> OF <input type="checkbox"/>							

This page intentionally left blank.

Appendix B: Statistical regression results

Effect of combat FitRep on PARS

Table 10 shows how each evaluation dimension differs on average when comparing combat and noncombat FitReps.

Table 10. Expected change in FitRep average (FRA) from combat FitRep, FY 1999 to FY 2003^a

	Evaluation dimension				
	Mission Perf.	Mission Prof.	Courage	Effectiveness	Initiative
Combat FitRep	0.21***	0.18***	0.48***	0.34***	0.07***
Grade controls	Yes	Yes	Yes	Yes	Yes
Year controls	Yes	Yes	Yes	Yes	Yes
Observations	48,937	48,805	43,877	45,109	48,458
	Leading	Developing	Set Example	Well-Being	Communication
Combat FitRep	0.04*	0.02	0.002	0.05***	0.04**
Grade controls	Yes	Yes	Yes	Yes	Yes
Year controls	Yes	Yes	Yes	Yes	Yes
Observations	44,910	43,877	48,603	44,125	48,605
	PME	Decisions	Judgment	Fulfill Evals.	
Combat FitRep	-0.10***	0.13***	0.11***	-0.02	
Grade controls	Yes	Yes	Yes	Yes	
Year controls	Yes	Yes	Yes	Yes	
Observations	47,772	48,157	48,166	27,032	

a. *** significant at the 1-percent level, ** significant at the 5-percent level, * significant at the 10-percent level.
Source: CNA analysis of Personnel Evaluation System data and Total Force Data Warehouse data.

Welcome-to-the-grade and room-to-grow effects

Table 11 shows the effect on FRAs and RO marks associated with a new grade or a new evaluator. Table 12 shows how these effects have changed in magnitude in recent years.

Table 11. Expected changes in FitRep average (FRA) and RO assessment, with separate effect for each MRO (“fixed effects”)^a

	All officers	2ndLt	1stLt	Capt	Maj	LtCol
Expected change in FRA						
Months commis- sioned	6.29 E-3***	0.03***	0.02***	5.74 E-3***	2.30 E-3**	-4.07 E-3**
First in grade	-0.06***	-0.03***	-0.02***	-0.03***	-1.00 E-2	0.01
First for RS	-2.01 E-3	0.02***	-0.02***	-0.02***	-3.33 E-3	0.01
First time with this RS (this grade)	-0.14***	-0.14***	-0.10***	-0.12***	-0.14***	-0.14***
Grade controls	Yes	N/A	N/A	N/A	N/A	N/A
Year controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	237,443	29,948	68,763	82,236	38,652	16,169
Unique MROs	23,850	13,454	17,630	16,482	8,983	4,786
Expected change in RO assessment						
Months commis- sioned	9.09 E-3***	0.04***	0.02***	0.01***	4.41 E-3**	7.21 E-4
First in grade	-0.09***	-0.02	-0.02	-2.00 E-2	-0.05***	-0.03
First for RO	0.14***	0.05***	0.06***	0.15***	0.19***	0.27***
First time with this RO (this grade)	-0.20***	-0.16***	-0.17***	-0.19***	-0.22***	-0.21***
Grade controls	Yes	N/A	N/A	N/A	N/A	N/A
Year controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	224,233	26,965	64,988	78,334	36,960	15,460
Unique MROs	23,620	12,856	17,472	16,235	8,797	4,614

a. *** significant at the 1-percent level, ** significant at the 5-percent level. Scientific notation represents multiplying or dividing by a power of 10. For example, 5.74 E-3 means 5.74 divided by 10^3 , or 0.00574.

Source: CNA analysis of Personnel Evaluation System data and Total Force Data Warehouse data.

Table 12. Expected change in FitRep average (FRA) or in RO assessment, with separate effect for each MRO (“fixed effects”) ^a

	FRA 1999–2005	FRA 2006–2011	RO mark 1999–2005	RO mark 2006–2011
Months commissioned	0.01***	4.40 E-3***	0.01***	8.22 E-3***
First in grade	-0.04***	-0.06***	-0.04***	-0.12***
First for RS/RO	0.04***	-0.06***	0.17***	0.11***
First time with this RS/RO (this grade)	-0.14***	-0.13***	-0.18***	-0.21***
Grade controls	Yes	Yes	Yes	Yes
Year controls	Yes	Yes	Yes	Yes
Observations	110,674	126,769	103,228	121,005
Unique MROs	14,722	20,374	14,525	20,192

a. *** significant at the 1-percent level.

Source: CNA analysis of Personnel Evaluation System data and Total Force Data Warehouse data.

Mutual predictive power of RV and RO mark

Table 13 shows both the tendency of RVs and RO marks to predict each other and their tendency to systematically differ.

Table 13. Expected changes in relative value (RV) and RO mark^a

	All officers		2ndLt	1stLt	Capt	Maj	LtCol
	Run 1	Run 2					
Expected change in RV							
RO mark this FitRep		2.04***	1.24***	1.88***	2.15***	2.21***	2.20***
RO mark cumulative average	0.28***	-1.45***	-0.98***	-1.42***	-1.36***	-1.33***	-1.62***
RV cum. avg.	0.91***	0.87***	0.96***	0.88***	0.84***	0.82***	0.79***
GCT	Yes	Yes	Yes	Yes	Yes	Yes	Yes
TBS 3 rd controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Commission source controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Grade controls	Yes	Yes	N/A	N/A	N/A	N/A	N/A
Year controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Marital status controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R-squared	0.481	0.539	0.831	0.591	0.492	0.394	0.346
Observations	166,971	159,464	15,966	45,677	59,507	26,350	11,000
Expected change in RO mark							
RV this FitRep		0.05***	0.04***	0.05***	0.06***	0.05***	0.05***
RV cum. avg.	0.01***	-0.03***	-0.02***	-0.02***	-0.03***	-0.02***	-0.02***
RO mark cumulative average	0.89***	0.85***	0.89***	0.90***	0.82***	0.81***	0.77***
GCT	Yes	Yes	Yes	Yes	Yes	Yes	Yes
TBS 3 rd controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Commission source controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Grade controls	Yes	Yes	N/A	N/A	N/A	N/A	N/A
Year controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Marital status controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R-squared	0.523	0.580	0.721	0.583	0.487	0.432	0.448
Observations	223,612	159,464	15,966	45,677	59,507	26350	11,000

a. *** significant at the 1-percent level, ** significant at the 5-percent level.

Source: CNA analysis of Personnel Evaluation System data and Total Force Data Warehouse data.

Quality

Table 14 shows the declining effect of TBS standing on FRA as officers progress in their careers. Table 15 compares the effect of different observable characteristics, including TBS standing, GCT score, commissioning source, and paygrade of the reporting senior.

Table 14. Expected change in FitRep average (FRA) from TBS third, by MRO grade^a

	All officers	2ndLt	1stLt	Capt	Maj	LtCol
Middle TBS third	-0.11***	-0.14***	-0.15***	-0.10***	-0.09***	-0.03**
Bottom TBS third	-0.17***	-0.25***	-0.25***	-0.14***	-0.12***	-0.05***
Gender control	Yes	Yes	Yes	Yes	Yes	Yes
Race/ethnicity controls	Yes	Yes	Yes	Yes	Yes	Yes
Marital status controls	Yes	Yes	Yes	Yes	Yes	Yes
Dependent controls	Yes	Yes	Yes	Yes	Yes	Yes
Grade controls	Yes	N/A	N/A	N/A	N/A	N/A
Year controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	164,317	16,992	39,536	61,440	33,729	12,620

a. *** significant at the 1-percent level, ** significant at the 5-percent level

Source: CNA analysis of Personnel Evaluation System data and Total Force Data Warehouse data.

Table 15. Expected change in FitRep average (FRA) from observable quality^a

	Run 1 (all officers)	Run 2 (all officers)	Run 3 (all officers)	Run 4 (all officers)	Run 5 (female officers)	Run 6 (male officers)	Run 7 (all officers)
Middle TBS third	-0.11***			-0.10***	-0.07***	-0.10***	-0.10***
Bottom TBS third	-0.17***			-0.16***	-0.15***	-0.16***	-0.17***
GCT score		-0.001***		-0.001***	-0.002***	-0.001***	-0.001***
Commissioning source							
ECP			Base	Base	Base	Base	Base
MCP			0.02	0.01	-0.02	0.01	0.01
OCC			-0.11***	-0.09***	-0.03	-0.09***	-0.06***
PLC			-0.14***	-0.12***	-0.02	-0.12***	-0.08***
MECEP			0.05***	0.02**	-0.06	0.03**	0.03***
NROTC			-0.03***	-0.03***	0.01	-0.03***	-0.01
USNA			-0.09***	-0.09***	-0.01	-0.09***	-0.05***
Other			0.02*	-0.05***	0.12**	-0.06***	-0.02**
RS Grade							
2ndLt							0.14
Capt							0.13
Maj							0.2
LtCol							0.53***
Col							0.83***
BGen							0.96***
MajGen							1.06***
LtGen							1.25***
Civilian							0.81***
Gender control	Yes	Yes	Yes	Yes	N/A	N/A	Yes
Race/ethnicity controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Marital status controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Dependent controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Grade controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	164,317	162,485	167,618	163,495	7,751	155,744	155,744

a. *** significant at the 1-percent level, ** significant at the 5-percent level, * significant at the 10-percent level.
Source: CNA analysis of Personnel Evaluation System data and Total Force Data Warehouse data.

Education

Table 16 shows the effects of academic qualifications on FitRep marks. Tables 17 and 18 compare the remaining predictive power of these qualifications when we account for the relationship between RVs and RO marks.

Table 16. Expected change in FitRep average (FRA) or RO mark among alumni of colleges with a Barron's competitiveness rating or the United States Naval Academy (USNA)^a

	FRA, ranked schools	RO, ranked school	FRA, USNA	RO, USNA
Most Competitive	0.05***	0.09***		
Highly Competitive	3.55 E-3	0.03		
Very Competitive	4.80 E-3	0.02		
Less Competitive	-0.02	-0.05		
Noncompetitive	-0.07*	-0.11 *		
GPA	0.09***	0.17***	0.10***	0.20***
Humanities	-5.47 E-3	-4.85 E-3	-0.01	8.26 E-3
Social Sciences	1.85 E-3	0.02	0.04**	0.03
Business	0.04***	0.04**	0.12***	0.13**
Science	-0.02 *	-7.78 E-3	0.01	-0.04
Engineering and Technology	0.02	0.04	-0.04*	-0.10***
Math	-0.01	0.04	0.05	-3.81 E-3
Professional	0.06**	0.11**	0.14	0.33*
Applied	-0.02	-1.56 E-3	-0.06	-0.24
Grade controls	Yes	Yes	Yes	Yes
Year controls	Yes	Yes	Yes	Yes
Occfield controls	Yes	Yes	Yes	Yes
R-squared	0.306	0.178	0.312	0.154
Observations	84,517	87,783	28,444	29,499

a. Errors clustered by MRO. Omitted categories: Competitive, unknown major.

"Special" schools are excluded.

*** significant at the 1-percent level, ** significant at the 5-percent level, * significant at the 10-percent level.

Source: CNA analysis of Personnel Evaluation System data and Total Force Data Warehouse data.

Table 17. Expected change in RV among alumni of colleges with a Barron's competitiveness rating

	Without occfield and major	With occfield and major
RO mark	3.11***	3.09***
Most Competitive or Highly Competitive	0.05	0.04
Less Competitive or Noncompetitive	-0.01	3.09 E-3
GPA	0.63***	0.63***
Grade controls	Yes	Yes
Year controls	Yes	Yes
Major controls		Yes
Occfield controls		Yes
R-squared	0.237	0.240
Observations	57,186	56,634

Table 18. Expected change in RO mark among alumni of colleges with a Barron's competitiveness rating^a

	Without occfield and major	With occfield and major
RV	0.07***	0.07***
Most Competitive or Highly Competitive	0.04***	0.03***
Less Competitive or Noncompetitive	-0.05*	-0.06**
GPA	0.10***	0.10***
Grade controls	Yes	Yes
Year controls	Yes	Yes
Major controls		Yes
Occfield controls		Yes
R-squared	0.342	0.345
Observations	57,186	56,634

a. Errors clustered by MRO. Omitted category: Very Competitive or Competitive. "Special" schools are excluded.

*** significant at the 1% level, ** significant at the 5% level,

* significant at the 10% level

Source: CNA analysis of Personnel Evaluation System data and Total Force Data Warehouse data

Race and ethnicity

Table 19 displays the gap between average FRAs or RO marks for black or Hispanic officers and white officers, and how these gaps change when we control for observable characteristics.

Table 19. Expected changes in FRA and RO mark^a

	Baseline	With GCT	Commiss. source	With academics	With TBS	TBS and occfield
Expected change in FRA						
Black	-0.06***	-0.06***	-0.07***	-0.05***	0.01	-0.06***
Hispanic	-0.03***	-0.03***	-0.04***	-0.04*	5.89 E-3	-0.03***
Grade controls	Yes	Yes	Yes	Yes	Yes	Yes
Year controls	Yes	Yes	Yes	Yes	Yes	Yes
GCT score		Yes				
Commiss. source controls			Yes			
School quality controls				Yes		
GPA				Yes		
Major controls				Yes		
TBS third controls					Yes	Yes
Occfield controls						Yes
R-squared	0.268	0.268	0.274	0.258	0.284	0.341
Observations	230,808	224,473	230,808	81,240	225,210	220,723
Expected change in RO mark						
Black	-0.17***	-0.16***	-0.19***	-0.18***	-0.04**	-0.10***
Hispanic	-0.10***	-0.10***	-0.12***	-0.13***	-0.03*	-0.06***
Grade controls	Yes	Yes	Yes	Yes	Yes	Yes
Year controls	Yes	Yes	Yes	Yes	Yes	Yes
GCT score		Yes				
Commiss. source controls			Yes			
School quality controls				Yes		
GPA				Yes		
Major controls				Yes		
TBS third controls					Yes	Yes
Occfield controls						Yes
R-squared	0.167	0.166	0.175	0.176	0.192	0.202
Observations	241,917	235,346	241,917	84,611	236,066	230,116

a. Errors clustered by MRO. Excluded category: white. Officers of “other” or “unknown” race/ethnicity are omitted.
 *** significant at the 1-percent level, ** significant at the 5-percent level, * significant at the 10-percent level.
 Source: CNA analysis of Personnel Evaluation System data and Total Force Data Warehouse data.

Gender match between RS and MRO

Table 20 shows the difference in behavior between male and female RSs.

Table 20. Expected change in FitRep Average (FRA) from RS gender^a

	Run 1	Run 2	Run 3
Female MRO	0.14***	0.13***	0.14***
Female RS		0.03***	0.05***
Female MRO*Female RS			-0.07***
Race/ethnicity controls	Yes	Yes	Yes
Marital status control	Yes	Yes	Yes
Dependent controls	Yes	Yes	Yes
TBS third controls	Yes	Yes	Yes
GCT control	Yes	Yes	Yes
Commissioning source controls	Yes	Yes	Yes
Grade controls	Yes	Yes	Yes
Year controls	Yes	Yes	Yes
Observations	163,495	156,245	156,245

a. *** significant at the 1-percent level.

Source: CNA analysis of Personnel Evaluation System data and Total Force Data Warehouse data.

Race/ethnicity match between RS and MRO

Table 21 shows the difference in behavior between white, black, Hispanic, and other RSs.

Table 21. Expected change in FitRep Average (FRA) from RS race/ethnicity^a

	White RS	Black RS	Hispanic RS	Other RS
MRO race/ethnicity				
White	Base	Base	Base	Base
Black	-0.02***	0.06***	-0.02	-0.03
Hispanic	0.002	0.01	0.02	-0.08**
Other	-0.03***	-0.01	0.03	0.02
Gender control	Yes	Yes	Yes	Yes
Marital status control	Yes	Yes	Yes	Yes
Dependent controls	Yes	Yes	Yes	Yes
TBS third controls	Yes	Yes	Yes	Yes
GCT control	Yes	Yes	Yes	Yes
Commissioning source controls	Yes	Yes	Yes	Yes
Grade controls	Yes	Yes	Yes	Yes
Year controls	Yes	Yes	Yes	Yes
Observations	136,617	6,859	7,311	3,940

a. *** significant at the 1-percent level, ** significant at the 5-percent level.

Source: CNA analysis of Personnel Evaluation System data and Total Force Data Warehouse data.

Occfield match

Tables 22 through 25 show how an RS or RO with the same occfield as the MRO changes the expected average level of FitRep marks in various paygrades and occfields. Table 26 shows how these effects change the spread, or variation, of FitRep marks.

Table 22. Expected change in RV, with separate effect for each MRO (“fixed effects”)^a

	All officers	2ndLt	1stLt	Capt	Maj	LtCol
Billet match	0.08	0.43	0.19	0.31**	-0.19	-0.73***
RS match	-0.23***	-0.80***	0.15*	-0.46***	-0.15	-0.40**
Exceptional billet	0.05	1.50*	-0.43*	0.40***	0.01	-0.43
Grade controls	Yes	N/A	N/A	N/A	N/A	N/A
Year controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	169,312	17,123	48,148	63,358	27,976	11,598
Unique MROs	23,395	10,100	16,028	15,238	7,684	3,883

a. *** significant at the 1-percent level, ** significant at the 5-percent level, * significant at the 10-percent level.
Source: CNA analysis of Personnel Evaluation System data and Total Force Data Warehouse data.

Table 23. Expected change in RV, with separate effect for each MRO (“fixed effects”)^a

	Personnel and admin	Intelligence	Logistics	Communications	Legal services
Billet match	0.32	0.43	-0.20	-0.02	0.33
RS match	0.24	-0.28*	0.26**	0.33*	-0.52**
Exceptional billet	-0.39	1.14***	-0.28	0.11	-0.33
Grade controls	Yes	Yes	Yes	Yes	Yes
Year controls	Yes	Yes	Yes	Yes	Yes
Observations	4,993	10,431	15,935	9,811	4,441
Unique MROs	742	1,872	2,335	1,589	803

a. *** significant at the 1-percent level, ** significant at the 5-percent level, * significant at the 10-percent level.
Source: CNA analysis of Personnel Evaluation System data and Total Force Data Warehouse data.

Table 24. Expected change in RO mark, with separate effect for each MRO (“fixed effects”)^a

	All officers	2ndLt	1stLt	Capt	Maj	LtCol
Billet match	0.08***	0.05	0.06***	0.05***	0.02	-0.04
RO match	-0.03***	-0.04*	9.27 E-3	-0.07***	-0.09***	-0.12***
RS match	-0.06***	5.59 E-3	-0.04***	-0.07***	-0.09***	-0.08***
Exceptional billet	0.02***	-0.43***	-0.09***	0.08***	0.10***	0.13***
Grade controls	Yes	N/A	N/A	N/A	N/A	N/A
Year controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	249,107	28,795	72,556	86,913	40,136	16,792
Unique MROs	24,655	13,381	18,186	17,159	9,256	4,852

a. *** significant at the 1-percent level, ** significant at the 5-percent level, * significant at the 10-percent level.
Source: CNA analysis of Personnel Evaluation System data and Total Force Data Warehouse data.

Table 25. Expected change in RO mark, with separate effect for each MRO (“fixed effects”)^a

	Personnel and admin	Intelligence	Logistics	Communications	Legal services
Billet match	0.27***	0.09***	0.10***	0.07*	0.15**
RO match	-0.46***	-0.06***	-0.13***	-0.06**	-0.23***
RS match	-0.10**	-0.19***	-0.10***	-0.03	-0.16***
Exceptional billet	0.02	-0.01	0.01	5.89 E-3	0.09
Grade controls	Yes	Yes	Yes	Yes	Yes
Year controls	Yes	Yes	Yes	Yes	Yes
Observations	7,478	17,722	23,674	14,967	7,009
Unique MROs	787	2,002	2,475	1,675	905

a. *** significant at the 1-percent level, ** significant at the 5-percent level, * significant at the 10-percent level.
Source: CNA analysis of Personnel Evaluation System data and Total Force Data Warehouse data.

Table 26. Expected change in square prediction errors^a

	RV	RO mark
Billet match	0.35	0.02
RO match	N/A	0.04***
RS match	-1.09***	-0.03***
Exceptional billet	2.67***	0.31***
Grade controls	Yes	Yes
Year controls	Yes	Yes
Observations	169,312	249,107

a. First, we predicted the RV or RO mark using billet match, RS and RO match, a control for exceptional billets, controls for paygrade and year, and fixed effects for each MRO. Then we regressed the squares of the residuals left over from this prediction on the same variables, but without an MRO fixed effect.
*** significant at the 1-percent level.

Source: CNA analysis of Personnel Evaluation System data and Total Force Data Warehouse data.

Promotion recommendation

Table 27 shows the effect of an officer's race or ethnicity and occupation on the wording of the RS's promotion recommendation while controlling for the RV generated by that RS's numerical marks.

Table 27. Expected change in promotion recommendation tier^a

	Including FitReps with RV of 80	Excluding FitReps with RV of 80
RV	-0.06***	-0.05***
Black	0.28**	0.24*
Hispanic	0.28**	0.23*
Ground	0.25**	0.32**
Aviation	-0.10	-0.07
Adjusted R-squared	0.167	0.135
Observations	283	255

a. Tier 1 is the strongest and 4 is the weakest, so lower numbers (a negative sign) indicate stronger recommendations

*** significant at the 1-percent level, ** significant at the 5-percent level, * significant at the 10-percent level.

Source: CNA analysis of Personnel Evaluation System data and Total Force Data Warehouse data.

Glossary

A-PES	Automated Performance Evaluation System
CY	Calendar Year
ECP	Enlisted Commissioning Program
E-to-O	Enlisted-to-Officer
EWS	Expeditionary Warfare School
FitRep	Fitness Report
FRA	FitRep Average
FY	Fiscal Year
GCT	General Classification Test
GPA	Grade Point Average
LDO	Limited Duty Officer
MCP	Meritorious Commissioning Program
MECEP	Marine Enlisted Commissioning Education Program
MM	Manpower Management Division
MMER	Manpower Management Evaluation Review
MMOA	Manpower Management Officer Assignment Branch
MMSB	Manpower Management Support Branch
MOS	Military Occupational Specialty
MRO	Marine Reported On
NROTC	Naval Reserve Officers Training Corps
OAD	Operations Analysis Division
OCC	Officer Candidates Course
Occfield	Occupational Field

PARS	Performance Anchored Rating Scales
PERB	Performance Evaluation Review Board
PES	Performance Evaluation System
PLC	Platoon Leaders Course
PME	Professional Military Education
RO	Reviewing Officer
RS	Reporting Senior
RV	Relative Value
TBS	The Basic School
TFDW	Total Force Data Warehouse
USNA	United States Naval Academy

List of figures

Figure 1.	RS FitRep averages (FRA) by rank of MRO and fiscal year	12
Figure 2.	Standard deviation of FitRep averages (FRA) by rank of MRO and fiscal year	13
Figure 3.	Average RO marks by rank of MRO and fiscal year.	14
Figure 4.	Distribution of RO marks among second lieutenant and lieutenant colonel FitReps compared with the intended RO mark distribution within a grade	15
Figure 5.	Average FitRep marks (FRA), by grade of MRO and years the RS has evaluated that grade	20
Figure 6.	Difference in average RV from peers in the same paygrade, year of report, and TBS third	43
Figure 7.	Difference in average RO mark from peers in the same paygrade, year of report, and TBS third	44

This page intentionally left blank.

List of tables

Table 1.	Performance anchored rating scales	7
Table 2.	Median FRA awarded by RS on each successive FitRep he or she writes for officers of a particular paygrade	22
Table 3.	Percentage of FitReps with an RV generated by the time the MRO is promoted, by number of reports and by number of different MROs needed to form RS profile	26
Table 4.	Distribution of RVs for FitReps with a particular RO mark among MRO captains, January 1999 to September 2011	30
Table 5.	Occfields with highest and lowest RVs and RO marks	36
Table 6.	Subjective comments classified into tiers of promotion recommendation strength	48
Table 7.	Strength of promotion recommendation for officers of each race or ethnicity	48
Table 8.	View of RV and RO mark tabulation in briefing guide for a fictional officer	55
Table 9.	Proposed alternative view of RV and RO mark tabulation in briefing guide	57
Table 10.	Expected change in FitRep average (FRA) from combat FitRep, FY 1999 to FY 2003.	73

Table 11.	Expected changes in FitRep average (FRA) and RO assessment, with separate effect for each MRO (“fixed effects”)	74
Table 12.	Expected change in FitRep average (FRA) or in RO assessment, with separate effect for each MRO (“fixed effects”)	75
Table 13.	Expected changes in relative value (RV) and RO mark	76
Table 14.	Expected change in FitRep average (FRA) from TBS third, by MRO grade.	77
Table 15.	Expected change in FitRep average (FRA) from observable quality	78
Table 16.	Expected change in FitRep average (FRA) or RO mark among alumni of colleges with a Barron’s competitiveness rating or the United States Naval Academy (USNA)	79
Table 17.	Expected change in RV among alumni of colleges with a Barron’s competitiveness rating.	80
Table 18.	Expected change in RO mark among alumni of colleges with a Barron’s competitiveness rating	80
Table 19.	Expected changes in FRA and RO mark	81
Table 20.	Expected change in FitRep Average (FRA) from RS gender	82
Table 21.	Expected change in FitRep Average (FRA) from RS race/ethnicity	83
Table 22.	Expected change in RV, with separate effect for each MRO (“fixed effects”)	84

Table 23.	Expected change in RV, with separate effect for each MRO (“fixed effects”)	84
Table 24.	Expected change in RO mark, with separate effect for each MRO (“fixed effects”)	84
Table 26.	Expected change in square prediction errors	85
Table 25.	Expected change in RO mark, with separate effect for each MRO (“fixed effects”)	85
Table 27.	Expected change in promotion recommendation tier	86

This page intentionally left blank.

